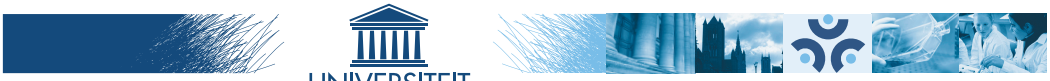

The Razor Boost analysis

Another step in the hunt for new physics at CMS

Nadja Strobbe

Promotoren: Prof. Dr. Dirk Ryckbosch, Dr. Michael Tytgat

Proefschrift ingediend tot het verkrijgen van de academische graad van Doctor in de Wetenschappen: Fysica

Vakgroep Fysica en Sterrenkunde, academiejaar 2014-2015

The important thing is not to stop questioning.
Curiosity has its own reason for existing.

Albert Einstein

Acknowledgements

Now that my time as a PhD student is coming to an end, it is time to reflect for a moment on the past 3.5 years. So much happened, so many people crossed my path. It all contributed to shaping the person I am today. I like to think it made me a better person. It definitely made me a more knowledgeable person. I learned so many new things, from physics and organization to just remembering to take the time to have fun every once in a while. Before the bravest among you embark on the effort to read through this thesis, I would like to give my thanks to everyone that helped me in any way throughout these past years.

First of all, I would like to thank my promotors Dirk and Michael, for allowing me to plunge into the adventure that is a PhD. Your doors were always open to answer my questions, or to give me advice. And above all, thank you for sending me off to CERN! That year-and-a-half was an experience I will never forget.

Sezen, without you this thesis could not have happened. I cannot thank you enough for adopting me as your student, and convincing me to join you in the razor boost effort. Working with you is always a pleasure, and I hope we can continue working on some joint projects in the future. Since Lukas introduced us way back in 2010, we have become great friends, and that is a gift to be treasured.

Harrison, every time we chat I feel like I have learned something. Your colourful way of expressing yourself never fails to bring a smile to my face. I am very honoured to have worked with you on this analysis, pioneering the way towards a proper statistical treatment.

I would also like to thank Maurizio, and the rest of the razor team, for providing help and insight whenever we were stuck.

Of course, a big thank you goes to all my friends. To the people of the INW, thank you for making it a fun workplace. To my friends at CERN, thank you for the crazy evenings at Meyrinoise, and the always entertaining Penn dinners. Bedankt aan al mijn krachtbalvrienden, de trainingen en wedstrijden die we samen meegemaakt hebben waren de ideale ontspanning tussen de drukte door. Karen en Mieke, met al onze drukke schema's hebben we elkaar niet zo vaak kunnen zien, maar ik wil dat jullie weten dat ik onze vriendschap koester. Ik hoop dat we in de toekomst ook in contact kunnen blijven!

Een dikke merci ook aan mijn ouders en broer. Jullie zorgden ervoor dat ik het contact met de 'echte wereld' niet verloor. Dankjewel voor alle steun

doorheen de jaren. Zonder jullie had ik dit niet gekund!

Brett, since I met you, my life has never been the same. We have seen and done so many fun things over the past couple of years. Thank you for being there through all the mad times, rushing to meet one or the other deadline. The last couple months were not easy, with us living 800 km apart. The evenings spent talking until past midnight about all our future plans really kept me going. Now, we are embarking on a brand new and super exciting adventure in Chicago. We cannot know what the future will hold, but I know that we can conquer anything together!

Contents

Chapter 1

Introduction

Curiosity. Among all human traits, curiosity is the one that really drives progress and innovation. Every child is born with it. It expresses itself as the relentless desire to learn about the world, be it by eating dirt because you wondered what it tastes like, or by going through your parents' cupboards even though that was not allowed. I am sure you can think of many more of these examples. To me, this fundamental human trait is what science is all about.

For centuries, scientists have investigated the world around them, pushing the boundaries of our knowledge and coming to an ever better understanding of nature. This exploration lead us to the biggest structures in the universe, the stars and galaxies, but also to the smallest. First, atoms were discovered, thought at the time to be indivisible. When Rutherford and Thomson showed that these atoms actually consist of a positive nucleus surrounded by negative charges, that was a paradigm-shifting revelation.

From then onward, theory and experiment worked hand-in-hand to come to the Standard Model of particle physics. Many new particles were discovered during the past one hundred years or so. The electron, proton and neutron were the first to be discovered. Then a range of particles, among which are the positron, pion and kaon, were identified in the study of cosmic rays. With the advent of particle accelerators, physicists were able to create and identify many more particles. There were so many new particles in fact, without any apparent structure to them, that they were called the 'particle zoo'. Only when Gell-Mann and Zweig proposed that these particles were composed of quarks, did pieces start falling into place. Experimental evidence of quarks was found not much later.

During the 1960s and 1970s, the theoretical formulation of the Standard Model took shape. The Standard Model describes the elementary particles and their interactions. In those days, physicists already knew of all particles belonging to the first two families: the electron and electron neutrino, the muon and muon neutrino, the up and down quarks, and the charm and strange quarks. The tau lepton and bottom quark were discovered in the second half of the 1970s, indicating the existence of a third family as well. If this family was to fit in the existing structure, two more particles needed to exist: the tau

neutrino, and the top quark. The Standard Model also predicted the existence of force carriers, in particular the gluon, and the W, Z and Higgs bosons. If the Standard Model was really to describe nature, then these particles needed to exist. And so the hunt for these new particles began.

The gluon was discovered indirectly in 1979 at DESY, and the W and Z bosons were observed in 1983 at CERN. The top quark was more difficult to find, it took until 1995 at Fermilab. The tau neutrino was observed directly for the first time in 2000, also at Fermilab. The only missing part was the elusive Higgs boson. In 2012, more than 40 years after it was first proposed, the Higgs boson was finally discovered at the CERN LHC.

The Standard Model has proven to be very accurate in its predictions, both in terms of particle content, and in behaviour of a multitude of processes. Until today, no significant deviations from the Standard Model predictions have been found. However, there are several indications that the current Standard Model is not the final word. Gravity is not included in the Standard Model for example, nor can it provide a dark matter candidate. There is also no good theoretical explanation for the mass of the Higgs boson.

In attempts to address these unresolved questions, many models of *new physics* have been proposed. One of the most popular models is supersymmetry (SUSY), which predicts the existence of even more new particles, namely a superpartner for every Standard Model particle. Many searches for signs of new physics have been performed already, first at LEP and Tevatron, and now at the LHC. Unfortunately, none of these searches have found any evidence. Nevertheless, the general feeling within the experimental community is still one of optimism. Run 2 of the LHC is just around the corner, opening up a thus far unexplored energy domain. No matter what will be found, our knowledge of the world around us will grow once more.

Over the past couple years I participated in two searches for new physics with the CMS experiment at the LHC. As a starting PhD student, I contributed to a search for the supersymmetric partner of the top quark, first using the dataset at 7 TeV centre-of-mass energy, and later also at 8 TeV. Once the first round of analyses was done, it had become clear that the standard searches had not found any evidence for new physics. Consequently, the focus within the search groups shifted towards more dedicated searches that aim to explore all possible gaps in the sensitivity. With the discovery of the Higgs boson, so-called natural supersymmetry also became a hot topic. The focus of these models is to provide a satisfying answer to the question of why the Higgs boson mass is relatively light, which is also known as the hierarchy problem.

In this thesis I present the *razor boost* analysis, which was my main research topic for the past two years. It especially targets natural SUSY models that were not well covered in previous analyses, and therefore fits perfectly within the scope of the late Run 1 analyses. The razor boost analysis uses the razor kinematic variables to search for signs of new physics in hadronic final states including a highly boosted W boson.

The thesis is structured in the following way: the first two chapters will briefly cover the Standard Model and the need for new physics beyond the Standard Model. An introduction to supersymmetry will be given in Chap-

ter 4, including natural supersymmetry and its phenomenological implications. Chapters 5.1 and 5.2 will provide some details on the LHC and the CMS experiment. The event generation, simulation, and reconstruction will be discussed in Chapter 6. Full details on the razor boost analysis are presented in Chapter 7, before concluding with a summary and outlook in Chapter 8.

Chapter 2

The Standard Model of particle physics

The Standard Model (SM) of particle physics encapsulates our current understanding of the elementary particles and their interactions. It was developed as a quantum field theory over the past fifty years and has been tested thoroughly by many different experiments. So far, the Standard Model has proven to be very effective at predicting and explaining a variety of physics processes. In this chapter I will give a short overview of some of the main characteristics of the Standard Model. For an in-depth and extensive discussion, I refer to Refs. [1–5]. Section 2.1 will cover the particle content and interactions present in the Standard Model. The phenomenon responsible for binding quarks together is discussed in Section 2.2, whereas the origin of the mass of the elementary particles is touched upon in Section 2.3. This chapter concludes with some of the biggest success stories of the Standard Model in Section 2.4.

2.1 Particles and interactions

The Standard Model contains three families (or generations) of fermionic matter, each consisting of a charged lepton, a corresponding neutrino, and an up- and down-type quark pair. The three charged leptons are the electron (e), muon (μ) and tau (τ). Their corresponding neutrinos are called the electron, muon, and tau neutrino (ν_e, ν_μ, ν_τ). The three quark families comprise the up (u) and down (d) quark, the charm (c) and strange (s) quark, and the top (t) and bottom (b) quark. Apart from the progressively larger mass of the particles, the three families are exact copies. Until this day, there is no explanation about why there are three families, and not more or less. In fact, the world around us is built solely from particles of the first generation. Each atom comprises a nucleus surrounded by electrons. The building blocks of the atomic nucleus are the protons and neutrons, which are combinations of up and down quarks (uud for the proton and udd for the neutron).

A theory containing only free particles would be quite uneventful. The

particles would simply exist. In Nature there are four fundamental interactions between particles that we currently know of: gravity, electromagnetism, the weak interaction, and the strong interaction.

The effects of gravity are seen all around us, even though it is by far the weakest of the four forces. It is the reason the apple falls to the ground, and the planets circle the Sun. Gravity is described by general relativity, and is a macroscopic theory that directly affects everything with a mass. No quantum theory of gravity has been developed yet, although much work has been done in that direction, and a force carrier, the graviton, has been proposed to exist. For the energy scales that are probed by particle collisions at the LHC, however, it can be safely ignored, being too weak to have an effect on the behaviour of the particles involved.

The remaining three interactions are the ones that are described by the Standard Model of particle physics. They can be formulated together, as a single, unified, quantum field theory, governed by the gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$. The interactions are mediated by the gauge bosons, which are vector (spin-1) particles. Particles can be charged under each part of this gauge group, meaning that they feel the corresponding interaction. An overview of the matter particles and their charges, the quantum numbers, is given in Table 2.1.

The gauge group $SU(3)_c$ describes quantumchromodynamics (QCD). This is commonly referred to as the strong interaction, the force that binds the quarks inside nucleons. QCD is mediated by the massless gluons, and affects particles that carry colour charge, of which there are three types: red (r), green (g) and blue (b). Within the Standard Model only the quarks and gluons carry colour charge. The strong force has thus no effect on leptons. Since the mediators themselves also carry colour charge, this leads to some interesting phenomena, as discussed in Section 2.2.

The remainder of the SM gauge group, $SU(2)_L \times U(1)_Y$ describes the electroweak interaction. The electroweak theory is a chiral theory, which means that left-handed and right-handed particles transform differently under the gauge group. This is also visible from Table 2.1. The charges of the electroweak interaction are the weak isospin T_3, and the hypercharge Y. The isospin reflects the chiral nature of the theory. Left-handed particles come in doublets, e.g. the u_L and d_L quarks, with weak isospin $\pm 1/2$. Right-handed particles, e.g. u_R, are singlets under $SU(2)_L$ and have weak isospin 0. The mediators are three massless W^i bosons and a massless B^0 boson, respectively. The W^i bosons form a weak isospin triplet, and the B^0 is a weak isospin singlet.

An important observation is that there cannot be explicit mass terms in the Lagrangian of a chiral theory, as this would break the gauge symmetry. All particles in the Standard Model should thus be massless. Of course, this is not what we observe in reality. We need a mechanism to introduce masses into our theory without spoiling gauge invariance. This mechanism is the Brout-Englert-Higgs mechanism, which is based on spontaneous symmetry breaking, and will be explained in more detail in Section 2.3. An important consequence of the symmetry breaking is that the W^3 and B^0 bosons mix to form the photon γ and the Z boson. The photon remains massless, an indication of the remaining $U(1)_{\mathrm{EM}}$ symmetry, but the W and Z bosons acquire mass. The

Fermions	Generation 1	2	3	Spin	Q	T_3	Y	Colour
Quarks	$\begin{pmatrix}\mathrm{u}\\ \mathrm{d}\end{pmatrix}_L$	$\begin{pmatrix}\mathrm{c}\\ \mathrm{s}\end{pmatrix}_L$	$\begin{pmatrix}\mathrm{t}\\ \mathrm{b}\end{pmatrix}_L$	$\frac{1}{2}$	$+\frac{2}{3}$ $-\frac{1}{3}$	$\frac{1}{2}$ $-\frac{1}{2}$	$\frac{1}{6}$	r,g,b
	u_R	c_R	t_R	$\frac{1}{2}$	$+\frac{2}{3}$	0	$\frac{2}{3}$	r,g,b
	d_R	s_R	b_R	$\frac{1}{2}$	$-\frac{1}{3}$	0	$-\frac{1}{3}$	r,g,b
Leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	$\frac{1}{2}$	0 -1	$\frac{1}{2}$ $-\frac{1}{2}$	$-\frac{1}{2}$	-
	e_R	μ_R	τ_R	$\frac{1}{2}$	-1	0	-1	-

Table 2.1: Overview of all fermions included in the Standard Model, along with the quantum numbers: electric charge Q, third component of weak isospin T_3, hypercharge Y, and colour.

conserved charge related to this $U(1)_{\mathrm{EM}}$ symmetry is the electric charge Q. We can express it in terms of weak isospin and hypercharge,

$$Q = T_3 + Y. \tag{2.1}$$

Electromagnetism can also be described by a quantum field theory: quantumelectrodynamics (QED). After electroweak symmetry breaking, the W^1 and W^2 are combined into the W^+ and W^- bosons, which now have a well-defined electric charge.

Because of this broken symmetry, we observe electromagnetism and the weak force as two distinct interactions in every day phenomena. Because the photon is massless, electromagnetism is a long-range interaction. It describes how electrically charged particles interact, and is the force that binds atoms and molecules together. The weak force is mediated by the massive W^+, W^- and Z bosons. Heisenberg uncertainty principle dictates that they can thus only exist for a short time. Therefore, the weak force, which governs radioactive decay, has only a short range.

2.2 QCD and asymptotic freedom

The key difference between QCD and QED is that the gluons interact with each other because they carry colour charge, unlike the photons, which carry no electric charge. The consequences of this difference become clear when we consider a given charge at different distance scales. For QED, a given charge appears smaller the further away the observer is. This property is called *screening*, and results from the polarization of the vacuum around the charge. For QCD the screening effect is also present for quark-antiquark pairs that are created in the vacuum. However, the gluons that appear also carry colour charge, but with a different polarizing effect. The further away one goes,

the more gluons are visible, and the larger the charge appears. Conversely, for very short distances, the *antiscreening* effect is much smaller, and the effective charge is reduced.

In general, the scaling of a coupling constant in a quantum field theory (α_S for QCD) is given by the renormalization group equations, and is called the β-function. For a non-Abelian quantum field theory like QCD, the β-function is given by

$$\beta(\alpha_S(\mu)) = \mu^2 \frac{d\alpha_S(\mu)}{d\mu^2} = -(b_0\alpha_S^2 + b_1\alpha_S^3 + \ldots), \tag{2.2}$$

with μ^2 the renormalization scale. The one-loop coefficient b_0 for QCD with n_f flavours of quarks is given by

$$b_0 = \frac{33 - 2n_f}{12\pi}. \tag{2.3}$$

The sign of b_0 determines whether the overall theory will exhibit screening or antiscreening. In our universe, only six quark flavours are known, resulting in a positive value for b_0, and an overall effect of antiscreening. It is also clear from Eq. 2.2 that a positive b_0 means that the size of the coupling constant diminishes as the momentum scale increases. This is what is called *asymptotic freedom*. The advantage is that quarks that are very close together can be treated as free quarks, and that perturbation theory holds. The validity of this picture was experimentally verified for the first time in deep-inelastic scattering. The current world average for the value of the strong coupling constant at the scale of Z boson mass is given by

$$\alpha_S(M_\mathrm{Z}) = 0.1185 \pm 0.0006, \tag{2.4}$$

and the running of the coupling is shown in Fig. 2.1.

At very low energy scales, the running of the strong coupling constant results in a very large value for α_S. This explains the observation that no one has ever found a single free quark. They are always 'confined' in colourless (white) hadrons, i.e. bound states of quarks and/or antiquarks. A quark and antiquark of opposite colour (e.g. r and $\overline{r}$) can form a *meson* and 3 quarks of different colour (r, g and b) can form a *baryon*. Mesons and baryons are collectively referred to as *hadrons*. When one tries to separate two quarks, for example in a high energy collision, they behave like a string, and energy is built up between them. At some point it becomes energetically more favourable to use this energy to create extra quarks from the vacuum. This process is called hadronization and is responsible for the creation of *jets*, the sprays of hadrons that are found at collider experiments. More information on hadronization is presented in Section 6.2.4.

2.3 Brout-Englert-Higgs mechanism

Introducing masses into a chiral theory is not a trivial issue, but is required to come to a viable description of the elementary particles. The Brout-Englert-Higgs mechanism [6, 7] accomplishes this by introducing one or more scalar

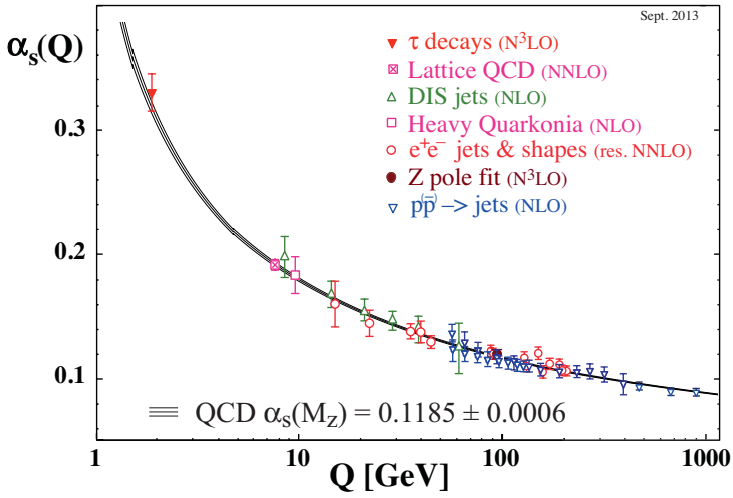

Figure 2.1: Latest summary of measurements of α_S as a function of the energy scale Q. The respective degree of QCD perturbation theory used in the extraction of α_S is indicated in parentheses. Figure and caption taken from Ref. [5].

fields, the Higgs fields, into the theory. These fields acquire a vacuum expectation value that spontaneously breaks a symmetry in the Lagrangian. According to the Goldstone theorem, every spontaneously broken continuous symmetry results in a massless scalar particle, the Goldstone boson. Hence, the number of Goldstone bosons in a theory is equal to the number of broken generators of the symmetry group. However, in the case of a gauge theory, like the Standard Model, this is not the full story. We shall see that the massless gauge bosons of the theory acquire a mass by absorbing the Goldstone bosons. The number of massive gauge bosons in a spontaneously broken gauge theory will thus be equal to the number of broken generators.

Before electroweak symmetry breaking, all four electroweak gauge bosons, W^1, W^2, W^3, and B^0, are massless. What we observe in experiment is one massless gauge boson, the photon γ, and three massive gauge bosons (W^+, W^-, Z). We also know that the electric charge is conserved. The spontaneous symmetry breaking should thus be of the form

$$SU(2)_L \times U(1)_Y \to U(1)_{\mathrm{EM}} \ .$$

For three gauge bosons to acquire mass, three Goldstone bosons will have to be absorbed. As a consequence, the scalar fields need to contain at least three degrees of freedom for the mechanism to work. The simplest way to do this is by introducing a complex, scalar $SU(2)$ doublet Φ with positive hypercharge $(Y = \frac{1}{2})$,

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \ . \tag{2.5}$$

The Standard Model Lagrangian without the strong part, i.e. only the electroweak gauge bosons and leptons, is given by

$$\mathcal{L}_{SM} = -\frac{1}{4}W^a_{\mu\nu}W_a^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \overline{L}_i(iD_\mu\gamma^\mu)L_i + \overline{e}_{R,i}(iD_\mu\gamma^\mu)e_{R,i} \ , \quad (2.6)$$

where i runs over the three generations, μ, ν are Lorentz indices and a runs over the number of generators in the gauge group. The field strengths are given by,

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g_2\epsilon^{abc}W^b_\mu W^c_\nu \ ,$$
$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \ ,$$

and the covariant derivative for left- and right-handed leptons by,

$$D_\mu L_L = \left(\partial_\mu - ig_2T_aW^a_\mu - ig_1YB_\mu\right)L_L \ ,$$
$$D_\mu e_R = \left(\partial_\mu - ig_1YB_\mu\right)e_R \ ,$$

where T_a are the generators of the $SU(2)_L$ gauge group and g_1, g_2 are the coupling constants for the electroweak interaction.

Having introduced the scalar doublet Φ, we need to add the corresponding scalar part to the Lagrangian,

$$\mathcal{L}_S = (D^\mu\Phi)^\dagger(D_\mu\Phi) - V(\Phi), \quad \text{with } V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 \ . \quad (2.7)$$

The first term is the kinetic term and the second term is the scalar potential, which is often called the "Mexican Hat" potential. The form of the Higgs potential is an assumption in the Standard Model. It is not known from first principles, but is rather chosen for its nice properties. It is the simplest potential that can achieve the necessary spontaneous symmetry breaking, while being renormalizable. In order for the vacuum to be stable, the parameter λ has to be positive. Depending on the sign of μ^2 one can distinguish two cases, which are illustrated in Fig. 2.2. In the case $\mu^2 > 0$, the potential $V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2$ is always positive with a the minimum at

$$\langle 0|\Phi|0\rangle \equiv \Phi_0 = \begin{pmatrix}0\\0\end{pmatrix} . \quad (2.8)$$

Since the minimum is at the origin, no spontaneous symmetry breaking takes place. In case $\mu^2 < 0$, the minimum of the potential is no longer located at the origin. Therefore, the neutral component of the scalar field can acquire a vacuum expectation value (vev) v, thereby breaking the electroweak symmetry.

$$\langle 0|\Phi|0\rangle \equiv \Phi_0 = \begin{pmatrix}0\\ \frac{v}{\sqrt{2}}\end{pmatrix} \ , \quad v = \sqrt{-\frac{\mu^2}{\lambda}} \ . \quad (2.9)$$

By only giving a vev to the neutral component, electromagnetism is conserved, as we set out to achieve.

We proceed by expanding Φ around its minimum Φ_0

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v + H(x)\end{pmatrix} \ , \quad (2.10)$$

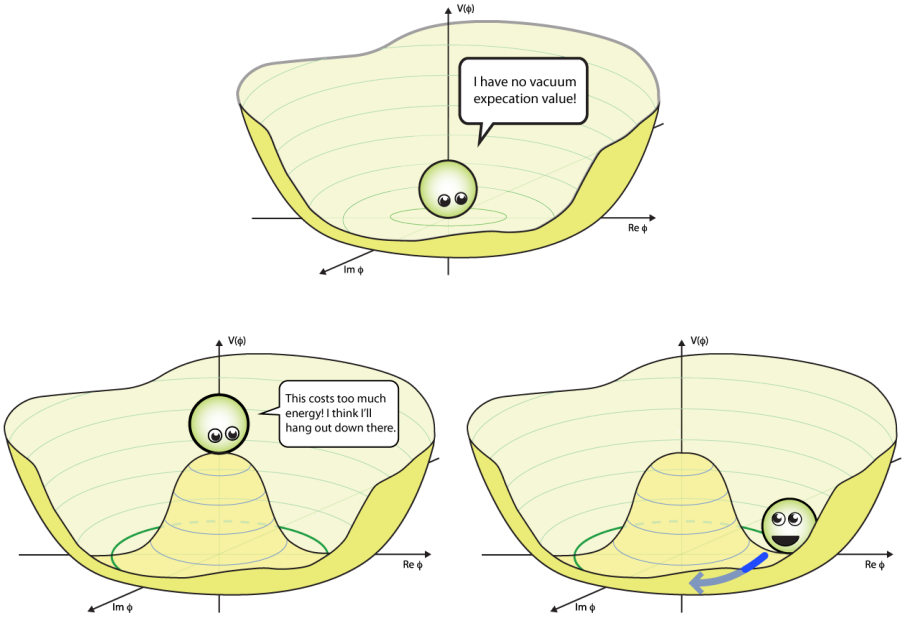

Figure 2.2: [top] The scalar potential with $\mu^2 > 0$ is always positive, with minimum in the origin. The scalar field will not obtain a vacuum expectation value. [bottom] The scalar potential with $\mu^2 < 0$ has a "Mexican hat" shape, with the minimum in a rim around the origin. The scalar field will thus move from the origin down to the actual minimum and acquire a vev in the process. The green line indicates a flat direction in the potential, corresponding to a massless Goldstone mode. Figures taken from Ref. [8].

where we have introduced a new scalar field $H(x)$. After inserting this in the kinetic part of the scalar Lagrangian (Eq. 2.7), and redefining the gauge fields as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \tag{2.11}$$

$$Z_\mu = \frac{1}{\sqrt{g_1^2+g_2^2}}(g_2 W_\mu^3 - g_1 B_\mu), \tag{2.12}$$

$$A_\mu = \frac{1}{\sqrt{g_1^2+g_2^2}}(g_2 W_\mu^3 + g_1 B_\mu), \tag{2.13}$$

we find for the kinetic part of the scalar Lagrangian,

$$|D_\mu \Phi|^2 = \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{2}g_2^2(v+H)^2 W_\mu^+ W^{\mu-} + \frac{1}{8}(v+H)^2(g_1^2+g_2^2)Z_\mu Z^\mu \ . \tag{2.14}$$

We see that the photon A_μ remains massless, as required for an unbroken symmetry. Mass terms for the W and Z bosons have the general form $M_W^2 W_\mu W^\mu$

and $\frac{1}{2}M_Z^2 Z_\mu Z^\mu$. We thus find for the gauge boson masses,

$$M_W = \frac{1}{2} v g_2, \tag{2.15}$$

$$M_Z = \frac{1}{2} v \sqrt{g_1^2 + g_2^2}, \tag{2.16}$$

$$M_A = 0\ . \tag{2.17}$$

After spontaneous symmetry breaking, three gauge bosons have thus absorbed a degree of freedom from the scalars (corresponding to the would-be Goldstone bosons), becoming massive in the process. One massless gauge boson and one scalar remain. The remaining scalar degree of freedom H corresponds to the so-called Higgs boson.

The mixing resulting in orthogonal combinations for the photon and Z bosons is often described in terms of the Weinberg or weak mixing angle, θ_W,

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3, \tag{2.18}$$

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3, \tag{2.19}$$

with the Weinberg angle itself defined as

$$\sin\theta_W = \frac{g_1}{\sqrt{g_1^2 + g_2^2}}. \tag{2.20}$$

At tree level, we also find a relation between the masses of the W and Z bosons:

$$\frac{M_W}{M_\mathrm{Z}} = \frac{g_2}{\sqrt{g_1^2 + g_2^2}} = \cos\theta_W. \tag{2.21}$$

The precise measurement of the W and Z boson masses can thus be used to measure the Weinberg angle. Working through the interaction terms between the photon and the fermions, one can show that the weak mixing angle relates the coupling strength of the weak interaction to that of the electromagnetic interaction,

$$e = g_2 \sin\theta_W. \tag{2.22}$$

The mass and couplings of the Higgs boson H can be determined from the scalar Lagrangian, Eq. 2.7, upon substituting Eq. 2.10. Using $v^2 = -\frac{\mu^2}{\lambda}$ and extracting the parts containing only H, we find for the Lagrangian of the Higgs boson:

$$\mathcal{L}_H = \frac{1}{2}(\partial_\mu H)(\partial^\mu H) - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4} H^4\ . \tag{2.23}$$

Scalar masses have the general form $\frac{1}{2}m\phi^2$; the Higgs boson mass is thus

$$m_H = 2\lambda v^2 = -2\mu^2, \tag{2.24}$$

and needs to be determined experimentally since there is no other way to access the parameter λ.

Now that we have generated masses for the gauge bosons, all we still need to do, is generate masses for the fermions as well. This can be done by introducing

Yukawa coupling terms between the fermions and the Higgs fields. The Yukawa Lagrangian for the first generation is given by

$$\mathcal{L}_F = -\lambda_e \overline{L}\Phi e_R - \lambda_d \overline{Q}\Phi d_R - \lambda_u \overline{Q}\widetilde{\Phi} u_R + h.c. \,, \tag{2.25}$$

where we introduced the conjugate of Φ, $\widetilde{\Phi} = i\tau_2\Phi^*$ which has negative hypercharge. This is needed to be able to couple to the up-type quarks. It is also possible to introduce a completely new Higgs doublet with negative hypercharge. This kind of model is called a two-Higgs doublet model (2HDM) and is needed to introduce supersymmetry (see Chapter 4).

Substituting (2.10) in (2.25), we find

$$\mathcal{L}_F = -\frac{1}{\sqrt{2}}\lambda_e(\overline{\nu}_e, \overline{e}_L)\begin{pmatrix} 0 \\ v+H \end{pmatrix} e_R + ... \tag{2.26}$$

$$= -\frac{1}{\sqrt{2}}\lambda_e(v+H)\overline{e}_L e_R + ... \,, \tag{2.27}$$

where we highlighted the electron part. Fermion mass terms have the general form $m_f \overline{f}_L f_R + h.c.$. We find

$$m_e = \frac{\lambda_e v}{\sqrt{2}}, \qquad m_u = \frac{\lambda_u v}{\sqrt{2}}, \qquad m_d = \frac{\lambda_d v}{\sqrt{2}}. \tag{2.28}$$

The neutrinos are seen to remain massless as a result of the lack of a right-handed neutrino in the Standard Model.

Using the same doublet of scalar fields we have thus successfully given mass to the gauge bosons and fermions in our theory. The Brout-Englert-Higgs mechanism also predicts the existence of a massive scalar particle, the Higgs boson. When in 2012 a particle with all the characteristics of this Higgs boson was found at the LHC, this meant the verification that the process of electroweak symmetry breaking is indeed realized in nature.

2.4 A success story

Since its conception decades ago, the Standard Model has performed beyond anyone's expectation. It has provided an accurate description of results from many accelerator and non-accelerator based experiments. In this section I will highlight the latest achievement, the discovery of the Higgs boson, and a global test of the validity of the Standard Model via electroweak precision fits.

2.4.1 Higgs boson discovery

The existence of the Higgs boson was first proposed in 1964 by Robert Brout and François Englert [6], and independently also by Peter Higgs [7]. Nearly half a century later, on the fourth of July, 2012, its discovery was finally announced by the CMS and ATLAS Collaborations at the LHC.

The allowed mass range for the Higgs boson had already been narrowed down by the experiments at LEP and Tevatron, and by the dataset delivered by the LHC in 2011 at a centre-of-mass energy of 7 TeV. In that latter

dataset, some evidence for a particle with a mass around 125 GeV was already observed [9], although not strong enough to claim a discovery. The energy increase to 8 TeV in 2012 provided just the boost needed to claim discovery with 5σ significance [10]. The evidence was strongest in the decay channels $H \to \gamma\gamma$ and $H \to \mathrm{ZZ}^* \to 4\ell$. Figure 2.3 shows the invariant mass distributions of the diphoton and four-lepton systems obtained by the CMS experiment. There is an excess visible around 125 GeV. The final Higgs boson mass measurement for Run 1 at CMS, combining all decay channels, and combining the 7 TeV and 8 TeV datasets, which correspond to an integrated luminosity of $5.1\,\mathrm{fb}^{-1}$ and $19.7\,\mathrm{fb}^{-1}$, respectively, found the following value for the Higgs boson mass,

$$m_H = 125.02\,^{+0.26}_{-0.27}\,(\mathrm{stat})\,^{+0.14}_{-0.15}\,(\mathrm{syst})\,\mathrm{GeV}. \tag{2.29}$$

The measured cross section σ also agrees very well with the expectation from the Standard Model. The signal strength at the measured mass was measured from the combination of all decay channels, and found to be

$$\frac{\sigma}{\sigma_{\mathrm{SM}}} = 1.00 \pm 0.09\,(\mathrm{stat})\,^{+0.08}_{-0.07}\,(\mathrm{theory}) \pm 0.07\,(\mathrm{syst}). \tag{2.30}$$

Many more tests of the properties of the observed particle have been made [11]. Examples are the spin and parity of the particle, the ratios of production rates for different production modes, the ratios of couplings to fermions and vector bosons, and the possible branching fraction to non-SM particles, $\mathrm{BR}_{\mathrm{BSM}}$. A subset of these measurements is shown in Fig. 2.4. The κ variables are the (effective) coupling strength modifiers w.r.t. the SM Higgs boson couplings. The couplings to massive vector bosons, fermions, gluons and photons are denoted by κ_{V}, κ_{f}, κ_{g} and κ_γ, respectively. The ratios of separate coupling strength modifiers are denoted by $\lambda_{\mathrm{XY}} = \kappa_{\mathrm{X}}/\kappa_{\mathrm{Y}}$. Figure 2.4 includes the ratio of couplings to W and Z bosons, λ_{WZ}, the ratio between down- and up-type quarks, λ_{du}, and the ratio of couplings between leptons and quarks, λ_{lq}. Excellent agreement with the Standard Model expectation is found.

2.4.2 Precision electroweak fits

Electroweak precision observables can probe energy scales much larger than what is possible in direct measurements, through the effects of higher-order corrections. Therefore, measurements of these observables at lepton colliders such as LEP and SLC, along with measurements at hadron colliders, such as Tevatron and LHC, need to be paired with very accurate theoretical predictions. Up until the Higgs boson discovery, this experimental and theoretical input was used in global fits of the electroweak sector in order to constrain the free parameters of the Standard Model. Now, on the other hand, assuming that the newly discovered Higgs boson is indeed the SM Higgs boson, all fundamental parameters that are used in the fit are known, and the fit is overconstrained. This means that we can now fully test the consistency of the Standard Model, as well as predict some of the observables with higher precision than the direct measurements. There are several groups performing these global fits [12–15]. I will take the latest results from the GFitter group [16, 17] as illustration here.

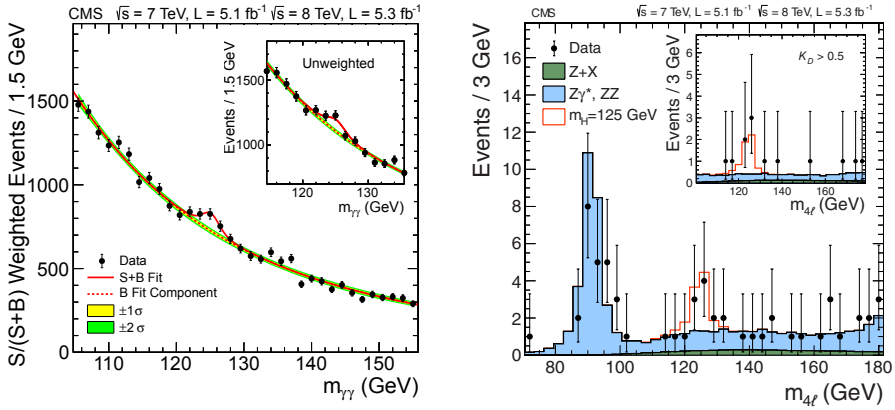

Figure 2.3: The Higgs discovery in the two high-resolution channels. Figures taken from Ref. [10]. [left] The diphoton invariant mass distribution with each event weighted by the $\frac{S}{S+B}$ value of its selection category. The lines represent the fitted background and signal, and the coloured bands represent the ± 1 and ± 2 standard deviation uncertainties in the background estimate. [right] Distribution of the four-lepton invariant mass for the $ZZ^* \to 4\ell$ analysis. The points represent the data, the filled histograms represent the background, and the open histogram shows the signal expectation for a Higgs boson of mass $m_H = 125\,\text{GeV}$, added to the background expectation.

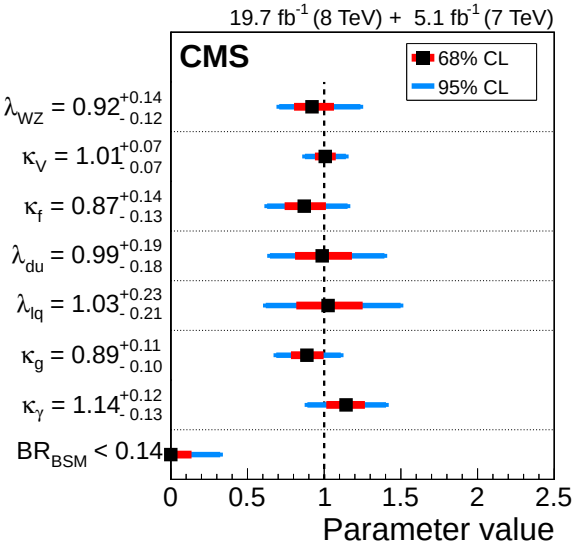

Figure 2.4: Summary of tests of the compatibility of the CMS data with the SM Higgs boson couplings as discussed in the text. The observed boson is fully compatible with the Standard Model expectation. Figure taken from Ref. [11].

The measurements that are used in the fit include, among others, the mass of the Higgs boson, the mass and width of the W and Z bosons, the mass of the top, bottom and charm quarks, the strong coupling constant, the forward-backward asymmetry and asymmetry parameters for leptons, c and b quarks, and the weak mixing angle. Several of the theoretical predictions that are used in the fit procedure are now known up to two or more loop orders. Examples are the effective weak mixing angle, the mass of the W boson, and the partial widths and branching ratios of the Z boson.

Figure 2.5 shows the comparison of the fit results with the direct measurements. The agreement is within three standard deviations for all observables. An important consistency test of the SM is the simultaneous indirect determination of the top and W boson mass. The top plot in Fig. 2.6 shows a scan of the confidence level profile of m_W versus m_{t} for the scenarios where the direct Higgs mass measurement is included in the fit (blue) or not (grey). Both contours agree with the direct measurements shown in the green bands and ellipses. The corresponding plot for the effective weak mixing angle and the W boson mass is shown on the bottom of Fig. 2.6. The agreement between the indirect determination via the fit procedure and the direct measurement, is very good, showing the consistency of the Standard Model with experiment.

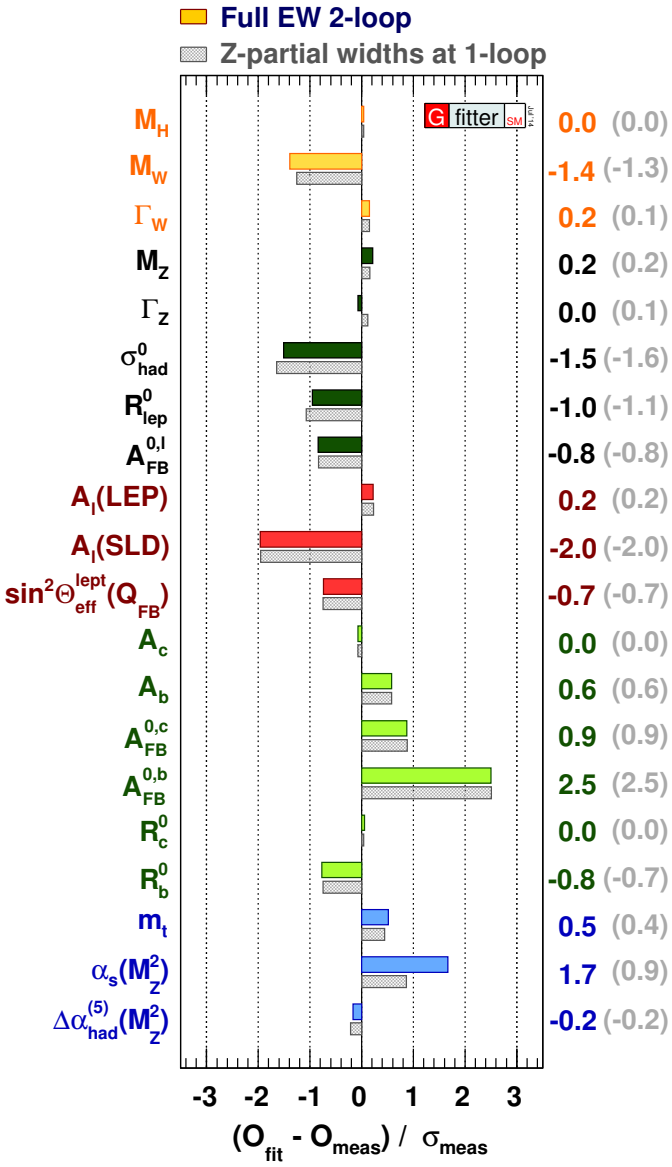

Figure 2.5: Comparison of the GFitter fit results with the direct measurements in units of the experimental uncertainty. Figure from Ref. [16].

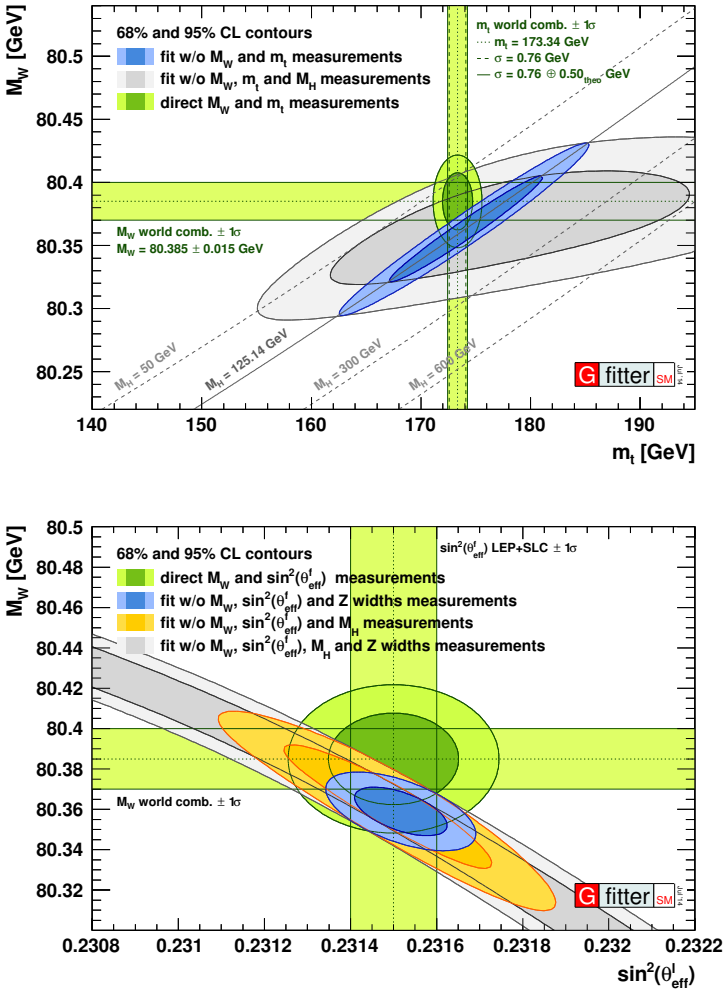

Figure 2.6: Contours at 68% and 95% confidence level obtained from scans of m_W versus m_{t} (top) and m_W versus $\sin^2(\theta^l_{\mathrm{eff}})$ (bottom), for a fit including m_H (blue) and excluding m_H (grey), as compared to the direct measurements (vertical and horizontal green bands and ellipses). In both figures, the corresponding direct measurements are excluded from the fit. Figures and caption from Ref. [16].

Chapter 3

The need for physics beyond the Standard Model

Although the Standard Model has succeeded in predicting and explaining a plethora of physics processes, it cannot be the ultimate theory describing Nature. In Section 3.1, I will list some open questions for which the Standard Model does not provide an answer. A very pertinent question regarding the Higgs boson mass is discussed in Section 3.2. Any new theory that attempts to solve these issues should reproduce the Standard Model for the energy domains that have been so thoroughly tested over the past decades. In this respect we can view the Standard Model as an effective low-energy theory, just like Newton's laws of motion follow from a low-energy approximation of special relativity. Most of the proposed new physics theories are, therefore, extensions of the Standard Model. A number of possibilities are introduced in Section 3.3.

3.1 Open questions

The open questions can roughly be divided in two categories. The first category comprises the experimental observations that are not covered at all by the SM, while the second category concerns characteristics of the SM for which we have no fundamental explanation. I provide a non-exhaustive list below.

Category 1

- **Gravity** is not included in the Standard Model. The main reason for this is that a satisfactory microscopic theory of gravity has not been formulated yet. Some advances have been made, e.g. supergravity theories [18–20], but it is not yet at the point where it can be unified with the rest of the Standard Model.

- The latest Planck results [21] state that about 26% of the energy budget of the universe is covered by **dark matter**, compared to less than about 5% for the ordinary matter that is described by the Standard Model. Currently, the Standard Model does not contain any particle that could be a dark matter candidate. Such a particle should be stable, neutral, weakly interacting and have a reasonably large mass.

- In addition to the presence of dark matter, we have very strong indications that the remaining 69% of the universe is **dark energy**, and drives the expansion of the universe. The Planck results also indicate that this dark energy is consistent with the assumption of a cosmological constant. The SM again does not provide an explanation for this, in fact, any attempt to compute this cosmological constant in terms of vacuum energy leads to a mismatch of around 100 orders of magnitude.

- The universe is almost entirely made up of matter, rather than antimatter. Assuming that equal amounts were created in the Big Bang, the SM cannot explain this **matter-antimatter asymmetry**. As postulated by Sakharov [22], there are three necessary conditions for a baryon asymmetry to exist: charge (C) and charge-parity (CP) violation; the absence of thermal equilibrium; and at least one baryon number violating process. Within the SM there is a small amount of CP violation, e.g. in the decay of the K^0 meson. However, even if the other two conditions were satisfied, there would not be enough CP violation to explain the observed matter-antimatter discrepancy.

- Neutrinos of different flavour have been observed to oscillate [23–25]. These **neutrino oscillations** can only occur if at least two of the three neutrino types have mass. The neutrino mass eigenstates (ν_1, ν_2, ν_3) are then superpositions of the flavour eigenstates $(\nu_e, \nu_\mu, \nu_\tau)$. No measurements of the absolute masses have been made so far, but the squared mass differences are known. In the Standard Model the neutrinos are massless. Adding a mass to the neutrinos can be done [26], but the question remains whether they are normal Dirac fermions, or Majorana fermions (i.e. their own antiparticle).

Category 2

- Why is the Higgs boson mass only 125 GeV? Radiative corrections would automatically drive this mass up to very large scales. This is often referred to as the hierarchy problem. As we will see later, this issue forms part of the motivation for the search presented in this thesis, and will thus be covered in some more detail in Section 3.2.

- Why are there three families of fermions? Careful study of the lineshape of the Z boson has shown that there is no fourth family with light neutrinos [27]. Could there be an extra family with heavier neutrinos?

- Why do the fermions have the masses, i.e. couplings to the Higgs boson, they have? And why is there such a wide range of masses, i.e. from

0.511 MeV for the electron up to 173 GeV for the top quark, a difference of 5 orders of magnitude. This is sometimes referred to as the fermion mass hierarchy problem.

- Are baryon and lepton numbers conserved? In the Standard Model these are accidental symmetries, without an underlying reason such as a local gauge symmetry. There is thus no compelling reason to assume that baryon and lepton number are conserved quantities.

- Why is the μ^2 parameter in the Higgs potential ($\mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2$) negative? Within the SM this is an assumption that is made, without underlying motivation other than that it is needed to trigger electroweak symmetry breaking.

3.2 The hierarchy problem

At the core of the hierarchy problem are the different mass scales that are present in the universe. The Standard Model does a very good job explaining phenomena at the electroweak scale of $\mathcal{O}(100\,\text{GeV})$. We know that when we reach the Planck scale, $\mathcal{O}(10^{19}\,\text{GeV})$, the SM can no longer be the complete theory, as quantum gravity effects will then need to be included.

Observed particle masses are a combination of the bare, tree-level mass, and all radiative corrections from additional loop diagrams. The loop momenta are cut off at the scale where we believe the theory to be no longer valid, in this case the Planck scale. For fermion masses these corrections are only logarithmically dependent on the high cutoff scale, as they are protected by chiral symmetry. Gauge bosons are similarly protected by the local gauge symmetry. The Higgs boson, being a scalar, does not have any protection, and therefore the radiative corrections introduce a quadratic dependence on the cutoff scale.

The one-loop corrections to the Higgs boson mass arise from the diagrams shown in Fig. 3.1. The fermionic loop correction arises from the Yukawa interaction between the Higgs boson and the fermions. The relevant part of the Lagrangian (Eq. 2.27) is given for a generic fermion as,

$$\mathcal{L}_{f\bar{f}H} = -\frac{\lambda_f}{\sqrt{2}} H f \bar{f}. \tag{3.1}$$

When one computes the fermionic one-loop diagram, we find for the correction to the Higgs boson mass [28],

$$\Delta m_H^2 = N_f \frac{\lambda_f^2}{8\pi^2}\left[-\Lambda^2 + 6m_f^2 \log\left(\frac{\Lambda}{m_f}\right) - 2m_f^2\right] + \mathcal{O}\left(\frac{1}{\Lambda^2}\right), \tag{3.2}$$

where the quadratic dependence on the cutoff scale Λ is explicitly visible. Taking this cutoff to be the Planck scale results in a correction that is more than 30 orders of magnitude larger than the Higgs boson mass squared itself. To still achieve an observable mass of 125 GeV, the bare mass and the correction would thus have to cancel to an extremely high precision. The inclusion of the vector

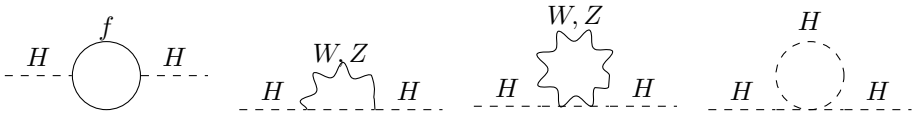

Figure 3.1: One-loop quantum corrections to the Higgs boson mass. From left to right: contribution from the Yukawa interaction; two contributions from the gauge interaction; contribution from the Higgs self-interaction.

boson and Higgs self-coupling loops do not change this overall behaviour. The requirement for this large amount of finetuning is viewed to be unnatural, and many models of new physics will use the *naturalness* argument in their favour.

A way to remove this quadratic dependence on the cutoff scale is to introduce extra particles in the theory, with properties such that the loop behaviour is opposite to the Standard Model particles. It is straightforward to show that we can cancel the fermionic loops by introducing extra scalar particles. Assuming that there are N_S new scalar particles, with mass m_S, trilinear coupling $v\lambda_S$ and quadrilinear coupling λ_S, we find as additional contribution to the one-loop correction to the Higgs mass:

$$\Delta m_H^2 = \frac{N_S \lambda_S}{16\pi^2}\left[-\Lambda^2 + 2m_S^2 \log\left(\frac{\Lambda}{m_S}\right)\right] - \frac{\lambda_S^2 N_S}{16\pi^2} v^2 \left[-1 + 2\log\left(\frac{\Lambda}{m_S}\right)\right] + \mathcal{O}\left(\frac{1}{\Lambda^2}\right). \quad (3.3)$$

By assuming $\lambda_f^2 = -\lambda_S$ and $N_S = 2N_f$, we find upon adding both contributions, and using Eq. 2.28,

$$\Delta m_H^2 = \frac{\lambda_f^2 N_f}{4\pi^2}\left[\left(m_f^2 - m_S^2\right)\log\left(\frac{\Lambda}{m_S}\right) + 3m_f^2 \log\left(\frac{m_S}{m_f}\right)\right]. \quad (3.4)$$

All quadratically divergent terms have vanished. Introducing scalar particles with the appropriate couplings has thus technically solved the hierarchy and naturalness problem. If in addition $m_S = m_f$, then the logarithmically divergent terms vanish as well.

The divergencies introduced by the other loop diagrams in Fig. 3.1 can also be resolved by the introduction of new particles, fermions in this case, that have just the right couplings to the Higgs boson. In this way all divergent contributions to the Higgs mass will vanish, and no large finetuning is needed.

3.3 Extensions of the Standard Model

In an attempt to address some of the afore-mentioned open questions, numerous models of beyond-the-SM (BSM) physics have been developed. With the discovery of the Higgs boson, some of these are now ruled out [29]. Examples are the Higgsless models such as the most basic incarnation of technicolour, or

the models which predict a very large Higgs boson mass, such as certain Composite Higgs models where the Higgs mass would be related to some new strong dynamics at high scales. Nevertheless, several viable models still remain. A subset of these are presented in the following sections.

3.3.1 Supersymmetry

Supersymmetric models impose a new symmetry, supersymmetry (SUSY), that relates fermions and bosons. Given that the razor boost analysis was developed with supersymmetry in mind and provides interpretations in a SUSY context, I will discuss supersymmetric models in a separate chapter (Chapter 4) and only provide a brief motivation here.

One of the nice features of SUSY is that it provides a solution to the hierarchy problem. The reason is that each known SM particle comes with a supersymmetric partner that differs in spin by 1/2, and has the same mass. The structure of the couplings is also exactly as we suggested in Section 3.2, resulting in the removal of the quadratically divergent terms in the mass correction. In practice we have not observed any such partner particles. Hence, their masses cannot be equal to the corresponding SM particles, and SUSY must be broken somehow. If the breaking mechanism is such that the quadratic divergencies still cancel, but not necessarily the logarithmic ones, then the hierarchy problem can still be solved.

SUSY also triggers electroweak symmetry breaking in a dynamic way. There is thus no need to explicitly assume a negative μ^2 in the Higgs potential. A very generic SUSY Lagrangian would allow interactions leading to proton decay. To avoid this, a new parity, called R-parity, is introduced (see also Section 4.3.3). If R-parity is conserved, then the lightest supersymmetric particle (LSP) is stable, as it cannot decay without violating R-parity. This LSP can be a dark matter candidate, as it would be heavy and only weakly interacting.

For all these reasons, and more, SUSY is nowadays by far the most popular extension of the Standard Model. Hopes are high to find hints of the existence of supersymmetric particles during the upcoming Run 2 of the LHC.

3.3.2 Little Higgs scenarios

In Little Higgs theories [29–31] the electroweak scale is stabilized in a natural way. The Higgs boson is viewed as a pseudo-Goldstone boson of a new global symmetry that is broken, both spontaneously and explicitly, by new physics around the 10 TeV scale. The Lagrangian contains two sets of interactions that explicitly break the symmetry, in addition to the symmetric part $\mathcal{L}_0$

$$\mathcal{L} = \mathcal{L}_0 + \lambda_1 \mathcal{L}_1 + \lambda_2 \mathcal{L}_2. \tag{3.5}$$

The Higgs boson would be an exact massless Goldstone boson if both couplings λ_1 and λ_2 vanish, and can only acquire a mass if both of them are present. This means that the corrections to the Higgs mass are suppressed by two loops w.r.t. the cutoff scale, and as such the hierarchy problem only appears around a scale of 10 TeV, compared to 1 TeV for the SM. Similarly to supersymmetry,

new particles are postulated to exist, but they should have the same spin as the known SM particles. Many choices for the new global symmetry can be made, resulting in slightly different model predictions. To avoid having a big impact on electroweak precision observables, T-parity is usually introduced. This parity ensures that the new particles have to be produced in pairs, as is the case for R-parity in SUSY, which means they only impact the observables at loop-level.

3.3.3 Extra dimensions

A key assumption in models of extra spatial dimensions [32], is that the electroweak scale is the only fundamental short distance scale. The loop corrections to the Higgs mass are thus cut off at the electroweak rather than the Planck scale, resulting in a much less finetuned model. The weakness of gravity is explained by assuming that gravity permeates these new dimensions, while the gauge interactions do not. The Planck scale in $(4 + n)$ dimensions is assumed to be of the order of the electroweak scale. The effective Planck scale at large distances (larger than the size R of the extra dimensions) then becomes $M_{\mathrm{Pl}(4+n)} \cdot R^n$. For $n \geq 2$, the needed size of the extra dimensions is sub-millimetre, a scale where the current understanding of gravity has not been tested yet. Models of extra dimensions can be tested in the high-energy collisions at the LHC [33]. One could detect excited gravitons, which preferentially decay to two high energy photons in some models, or to two quarks in other models. Another possibility is to look for missing energy when particles disappear into the extra dimensions.

Chapter 4

Supersymmetry

Supersymmetry is a space-time symmetry that relates fermions and bosons. The operator Q generating the transformation from fermion to boson and vice versa is an anti-commuting spinor, with

$$Q\,|\text{Boson}\rangle = |\text{Fermion}\rangle \tag{4.1}$$
$$Q\,|\text{Fermion}\rangle = |\text{Boson}\rangle\,. \tag{4.2}$$

From the above we can immediately see that Q must be fermionic in nature, and carries spin 1/2. The generator Q and its hermitian conjugate $Q^\dagger$ satisfy the anticommutation and commutation relations

$$\{Q, Q^\dagger\} = P^\mu, \tag{4.3}$$
$$\{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0, \tag{4.4}$$
$$[P^\mu, Q] = \left[P^\mu, Q^\dagger\right] = 0, \tag{4.5}$$

with P^μ the four-momentum generator of spacetime translations. Particles in a supersymmetric theory fall in the irreducible representations of the SUSY algebra, the supermultiplets, which contain both a fermion and a boson. Since the SUSY generators commute with P^2, it follows that the particles within a given supermultiplet have the same mass. Since we have not experimentally observed any of these new particles, this implies that SUSY must be a broken symmetry. Given that Q and $Q^\dagger$ also commute with the generators of the gauge transformations, particles in the same supermultiplet must have the same gauge quantum numbers.

Supersymmetry was originally introduced [34–40] to help construct a grand unified theory (GUT), merging all known forces in a single framework. For this to occur, there needed to be a link between particles of different spin. In this context it is also noteworthy to mention that the coupling constants appear to unify at around 10^{16} GeV if we include contributions from supersymmetry in the renormalization group equations. Making SUSY a local symmetry results into supergravity, which is the first step to including gravity in the overall unified framework.

Nowadays, supersymmetry receives much attention due to the solution it can provide for the hierarchy problem, and the dark matter candidate that is present in many SUSY models. If SUSY is to be the solution to the hierarchy problem, then this also means that at least some of the supersymmetric partners (superpartners) of the ordinary SM particles should be produced at the LHC energies. This is what motivates the many SUSY searches performed at the CMS and ATLAS experiments, including the search presented in this thesis.

In this chapter I will first provide some general considerations on possible supermultiplets and SUSY breaking mechanisms in Sections 4.1 and 4.2. Then, I will focus on the so-called Minimal Supersymmetric Standard Model (MSSM) in Section 4.3, covering its particle content, and introducing the concept of R-parity. Section 4.4 will explain how to break the electroweak symmetry in a supersymmetry context. The production and decay of superpartners along with their basic experimental signatures are presented in Section 4.5. The last two sections cover topics that are of particular importance to the razor boost analysis. Section 4.6 will cover natural SUSY and how it will solve the hierarchy problem. The concluding section in this chapter will cover the simplified approach that is used by experiments to guide new physics searches and present results. The bulk of this chapter is based on the excellent review by S. Martin [41].

4.1 Supermultiplets

Particles in a supersymmetric theory are represented by the supermultiplets, also called superfields. These are spin multiplets which contain both fermion and boson states. From the spin-statistics theorem it follows that there must be an equal number of fermionic and bosonic degrees of freedom. Interactions between particles can be most elegantly formulated in terms of the superfields, rather than the separate components. Superfields live in superspace, an extension of the usual spacetime with fermionic (Grassmann) variables. A detailed discussion of the superfield formalism is, however, beyond the scope of this thesis. Interested readers are referred to Refs. [41, 42].

The simplest possibility for a supermultiplet consists of a two-component Weyl fermion and a complex scalar field, the sfermion. It is called a *chiral* or *matter supermultiplet*. The ordinary SM fermions fit into these chiral multiplets, albeit that the left- and right-handed fermions need to reside in different multiplets.

A second possibility is to include a spin-1 vector boson. The resulting supermultiplet is called a *gauge* or *vector supermultiplet* and consists of a massless spin-1 boson and a massless spin-1/2 Weyl fermion. The bosons can only attain a mass through spontaneous symmetry breaking if the theory is to remain renormalizable. Given that gauge bosons transform under the adjoint representation of the gauge group, so must their superpartners, the so-called gauginos. This means that left-handed and right-handed gauginos will transform the same under the gauge group as opposed to the fermions we know from the Standard Model.

There are other possibilities to construct supermultiplets, but if they have

renormalizable interactions they can all be reduced to chiral and gauge supermultiplets. Hence, these will not be considered here. How all SM particles fit inside the supermultiplets will be discussed in Section 4.3.1, when introducing the particle content of the MSSM.

4.2 Supersymmetry breaking mechanisms

Unbroken supersymmetry leads to superpartners with the same mass as the normal SM particles, and which would thus have been discovered a long time ago. Since this is not the case, supersymmetry, assuming it exists, must be broken such that the superpartners have a large mass and would have avoided detection.

As we will see shortly, this has an influence on the discussion of the hierarchy problem in a SUSY context. When assuming perfect supersymmetry, there is no correction to the Higgs boson mass at all. When supersymmetry is broken, on the other hand, this cancellation no longer happens. In order to avoid the reappearance of quadratic divergencies, we can only allow *soft SUSY breaking*. This means that the relations between the dimensionless coupling constants (e.g. $\lambda_S = |\lambda_f|^2$, cf. Section 3.2) must still hold. If this is the case we will only have logarithmically divergent terms contributing to the Higgs mass correction, which will be of the form

$$\Delta m_H^2 = m_{\text{soft}}^2 \left(\frac{\lambda}{16\pi^2} \ln \frac{\Lambda}{m_{\text{soft}}} + \dots \right), \tag{4.6}$$

with m_{soft} the mass scale associated with the soft terms, λ a schematic representation of various couplings and where the ellipses include higher order loop corrections and terms independent of the ultraviolet cutoff scale Λ. Therefore, the masses of the superpartners cannot be too huge, otherwise the m_{soft}^2 corrections to the Higgs mass would become unnaturally large again.

Unfortunately, breaking SUSY softly is not so easy, and, up to now, nobody has found a satisfactory way to do it dynamically. There are several ideas, which are mostly based on the idea that SUSY is broken in a different sector, and then communicated to the visible sector by some form of mediator particles. The two main classes of mediator mechanisms are via gravitational interactions, or via the gauge interactions.

Because of these difficulties, the possible soft SUSY breaking terms are usually added to the Lagrangian by hand,

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{soft}}, \tag{4.7}$$

resulting in a low-energy effective theory. This effective theory can then be used to predict masses, decays etcetera. The downside is of course that the parameters governing the soft breaking terms are mostly unconstrained.

4.3 Minimal Supersymmetric Standard Model

4.3.1 Particle content

A first step to building a minimal supersymmetric extension of the Standard Model, is to fit all the SM particles in supermultiplets. Standard Model fermions – the quarks and leptons – have to be members of chiral supermultiplets, because left- and right-handed fermions transform differently under the electroweak gauge symmetry. Their spin-0 superpartners are called squarks and sleptons, where 's' stands for scalar. Symbolically, superpartners are denoted with a tilde above the usual SM symbol. Often a 'handedness' is also assigned to the superpartners, but it is important to remember that they are scalars, and so the concept of helicity is ill-defined. The handedness in this case is just a label referring to their SM partners.

The Higgs fields are spin-0 fields, and so they will also reside in a chiral supermultiplet. The fermionic partner of the Higgs field is called a higgsino. As it turns out, one Higgs chiral supermultiplet is not enough in supersymmetry. Two such multiplets are needed for two main reasons.

1. Anomaly cancellation: an anomaly occurs when a classical symmetry of the Lagrangian is not conserved at the quantum level. If this happens for a gauge symmetry, this causes the theory to be inconsistent. The specific particle content of the SM results in an anomaly free gauge symmetry because the condition $\mathrm{Tr}[T_3^2 Y] = \mathrm{Tr}[Y^3] = 0$, with the trace running over the left-handed Weyl fermionic degrees of freedom in the theory, is satisfied. A higgsino must have either hypercharge $Y = +1/2$, or $Y = -1/2$. By itself, the introduction of a higgsino would thus spoil the nice cancellation in the trace. To restore it, we need two Higgs supermultiplets with opposite hypercharge.

2. Structure of supersymmetric theories: only a $Y = +1/2$ Higgs chiral supermultiplet can have the Yukawa couplings needed to give masses to up-type quarks and, similarly, only a $Y = -1/2$ Higgs can give mass to down-type quarks and charged leptons. In the SM, the conjugate of the Higgs field is used to give mass to the down-type quarks, but this is no longer possible in SUSY because the conjugate would have the wrong chirality.

The two Higgs supermultiplets will be denoted by H_u and H_d, for $Y = +1/2$ and $Y = -1/2$, respectively.

The chiral supermultiplets present in the MSSM are summarized in Table 4.1. The representation of the Standard Model gauge group under which the supermultiplets transform is given in the last column. All chiral supermultiplets are defined in terms of left-handed Weyl spinors, which is why the conjugates of the right-handed quarks and leptons appear in the table.

The vector bosons of the Standard Model will have to reside in gauge supermultiplets together with a Majorana fermion field. An overview of the gauge superfields present in the MSSM is given in Table 4.2. After electroweak symmetry breaking, the neutral gauge bosons W^3 and B^0 mix to form the mass

Table 4.1: Chiral supermultiplets in the MSSM with their gauge quantum numbers.

Names		Spin 0	Spin 1/2	$SU(3)_c \times SU(2)_L \times U(1)_Y$
squarks, quarks (3 families)	$\widehat{Q}$	$(\widetilde{\mathrm{u}}_L\ \widetilde{\mathrm{d}}_L)$	$(\mathrm{u}_L\ \mathrm{d}_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	$\widehat{U}^c$	$\widetilde{\mathrm{u}}_R^*$	$\mathrm{u}_R^\dagger$	$(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	$\widehat{D}^c$	$\widetilde{\mathrm{d}}_R^*$	$\mathrm{d}_R^\dagger$	$(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons (3 families)	$\widehat{L}$	$(\widetilde{\nu}\ \widetilde{e}_L)$	$(\nu\ e_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	$\widehat{E}^c$	$\widetilde{e}_R^*$	$e_R^\dagger$	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, higgsinos	$\widehat{H}_u$	$(H_u^+\ H_u^0)$	$(\widetilde{H}_u^+\ \widetilde{H}_u^0)$	$(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$
	$\widehat{H}_d$	$(H_d^0\ H_d^-)$	$(\widetilde{H}_d^0\ \widetilde{H}_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$

Table 4.2: Gauge supermultiplets in the MSSM with gauge quantum numbers.

Names		Spin 0	Spin 1/2	$SU(3)_c \times SU(2)_L \times U(1)_Y$
gluino, gluon	$\widehat{G}_a$	$\widetilde{g}_a$	g^μ	$(\mathbf{8}, \mathbf{1}, 0)$
winos, W bosons	$\widehat{W}_a$	$\widetilde{W}_a$	W_a^μ	$(\mathbf{1}, \mathbf{3}, 0)$
bino, B^0 boson	$\widehat{B}$	$\widetilde{B}$	B^μ	$(\mathbf{1}, \mathbf{1}, 0)$

eigenstates Z and γ. The corresponding gaugino mixed states are called the zino ($\widetilde{\mathrm{Z}}$) and photino ($\widetilde{\gamma}$).

It is interesting to observe that mass terms for the sfermions, higgsinos and gauginos are not strictly forbidden. Squarks and sleptons are scalar fields, and a mass term $m^2|\phi|^2$ is always allowed by gauge symmetries. For the higgsinos and gauginos this is allowed because they are fermions in a real representation of the gauge group. The known SM particles on the other hand, would all be massless if not for the electroweak symmetry breaking. Therefore, from the point of view of the MSSM, it does not come as a surprise that those are the only particles that were light enough to be detected so far.

4.3.2 Supersymmetric Lagrangian

A Lagrangian for any interacting theory consists of kinetic terms and interaction terms. The full derivation of the supersymmetric Lagrangian will not be given here. Several extra concepts beyond the scope of this thesis would need to be introduced. That said, I will still provide the final components of the Lagrangian in this section, as they will be used in Section 4.4 in the discussion of electroweak symmetry breaking.

The kinetic terms for the MSSM are the supersymmetric equivalents of those of the SM (written as two-component spinors), and are given by

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}G^a_{\mu\nu}G_a^{\mu\nu} + \widetilde{G}^{\dagger a}i\overline{\sigma}^\mu D_\mu \widetilde{G}_a + f^{\dagger i}i\overline{\sigma}^\mu D_\mu f_i - (D_\mu\phi_i)^\dagger D^\mu\phi^i, \quad (4.8)$$

with G shorthand for any gauge boson, $\widetilde{G}$ for any gaugino, f for any chiral fermion and ϕ for any scalar.

The interaction part of any supersymmetric Lagrangian for the chiral superfields can be obtained using the superfield formalism from a superpotential $\mathcal{W}$,

$$\mathcal{L}_{\text{int}} = -\sum_i \left|\frac{\partial\mathcal{W}}{\partial z_i}\right|^2 - \frac{1}{2}\sum_{ij}\left(\overline{f}_i\frac{\partial^2\mathcal{W}}{\partial z_i\partial z_j}f_j + h.c.\right), \quad (4.9)$$

where z_i are the superfields in the theory. The particular superpotential giving rise to the supersymmetric interaction Lagrangian of the MSSM is given by

$$\mathcal{W} = \sum_{i,j} -Y^u_{ij}\widehat{U}^c_i\widehat{H}_u\cdot\widehat{Q}_j + Y^d_{ij}\widehat{D}^c_i\widehat{H}_d\cdot\widehat{Q}_j + Y^l_{ij}\widehat{E}^c_i\widehat{H}_d\cdot\widehat{L}_j + \mu\widehat{H}_u\cdot\widehat{H}_d, \quad (4.10)$$

where i,j run over the three generations and Y_{ij} represents the Yukawa coupling among generations. The first three terms are direct generalizations of the Yukawa interactions present in the Standard Model. The last term is new, and is a globally supersymmetric mass term for the two Higgs fields.

In addition to the interactions included in $\mathcal{L}_{\text{int}}$, there are also extra interaction terms involving both sfermions and gauginos:

$$\sqrt{2}g\widetilde{G}\phi^*T_af + h.c. + \frac{g^2}{2}\left|\phi^*T_a\phi\right|^2. \quad (4.11)$$

These terms are the supersymmetric counterparts of the ordinary fermion-gauge boson couplings.

4.3.3 R-parity

Apart from the interaction terms listed in the previous section, the most general MSSM Lagrangian could also include lepton and baryon number violating interactions such as,

$$W_{\Delta L=1} = \frac{1}{2}\lambda^{ijk}L_iL_j\bar{e}_k + \lambda'^{ijk}L_iQ_j\bar{\mathrm{d}}_k + \mu'^iL_iH_u, \quad (4.12)$$

$$W_{\Delta B=1} = \frac{1}{2}\lambda''^{ijk}\bar{\mathrm{u}}_i\bar{\mathrm{d}}_j\bar{\mathrm{d}}_k, \quad (4.13)$$

$$(4.14)$$

where i,j,k indicate family indices. These interactions could result in proton decay, $p\to\pi^0+e^+$. Experimentally we know that protons have a lifetime $>10^{32}$ years, such that these types of interactions should be suppressed. Therefore, a discrete, multiplicative symmetry, called R-parity, is imposed. Its quantum number is given by:

$$R = (-1)^{3B+L+2s}, \quad (4.15)$$

with B, L and s the baryon number, lepton number and spin quantum number, respectively. Regular Standard Model particles have $R = 1$, their supersymmetric partners have $R = -1$. Only terms conserving R-parity are allowed in the supersymmetric Lagrangian.

The introduction of R-parity has important consequences for the phenomenology of the MSSM. At particle colliders which collide Standard Model particles, supersymmetric particles can only be produced in pairs. Once produced, a SUSY particle will always decay into an odd number of other SUSY particles. The lightest supersymmetric particle (LSP) will therefore necessarily be stable, and can be a dark matter candidate if it is neutral and thus only weakly interacting.

4.3.4 Soft SUSY breaking in the MSSM

The MSSM Lagrangian follows the general form of Eq. 4.7, containing both a supersymmetry conserving part, and a supersymmetry breaking part. The SUSY conserving part is the direct translation of the SM Lagrangian using superfields rather than the SM fields, and was given in Section 4.3.2. The soft SUSY-breaking terms in the MSSM are mainly related to the superpartner masses. As explained before, gauge invariance does not restrict the appearance of their mass terms. All allowed soft SUSY breaking terms in the Lagrangian are:

1. gaugino mass terms:

$$-\mathcal{L}_{\text{gaugino}} = \frac{1}{2}\left(M_1\widetilde{B}\widetilde{B} + M_2\sum_{a=1}^{3}\widetilde{W}^a\widetilde{W}_a + M_3\sum_{a=1}^{8}\widetilde{g}^a\widetilde{g}_a + h.c.\right) \quad (4.16)$$

2. scalar fermion mass terms:

$$-\mathcal{L}_{\text{sfermion}} = \sum_{i=1}^{3}\Big(m^2_{\widetilde{Q}_i}\widetilde{Q}_i^\dagger\widetilde{Q}_i + m^2_{\widetilde{L}_i}\widetilde{L}_i^\dagger\widetilde{L}_i + m^2_{\widetilde{e}_{Ri}}\left|\widetilde{e}_{Ri}\right|^2 + m^2_{\widetilde{u}_{Ri}}\left|\widetilde{u}_{Ri}\right|^2 + m^2_{\widetilde{d}_{Ri}}\left|\widetilde{d}_{Ri}\right|^2\Big) \quad (4.17)$$

3. mass and bilinear terms for the Higgs bosons:

$$-\mathcal{L}_{\text{Higgs}} = m^2_{H_u}H_u^\dagger H_u + m^2_{H_d}H_d^\dagger H_d + B\mu(H_u\cdot H_d + h.c.) \quad (4.18)$$

4. trilinear couplings between sfermions and Higgs bosons:

$$-\mathcal{L}_{\text{trilinear}} = \sum_{i,j=1}^{3}\Big(A^u_{ij}Y^u_{ij}\widetilde{u}^*_{Ri}H_u\cdot\widetilde{Q}_j + A^d_{ij}Y^d_{ij}\widetilde{d}^*_{Ri}H_d\cdot\widetilde{Q}_j + A^l_{ij}Y^l_{ij}\widetilde{e}^*_{Ri}H_d\cdot\widetilde{L}_j + h.c.\Big) \quad (4.19)$$

These SUSY breaking terms introduce a large amount of free parameters in the theory. In total there are 105 extra parameters (masses, phases, mixing angles) compared to the Standard Model. Some of these new parameters can induce extra flavour mixing or CP violation, for which no signs have been seen experimentally. Therefore, one often assumes that SUSY breaking is flavour-blind and does not introduce new complex phases, which results in a model with far fewer parameters.

4.4 Electroweak symmetry breaking and natural SUSY

In this section I will explain how SUSY breaking automatically triggers electroweak symmetry breaking, thereby removing the ad hoc assumption of a negative μ^2 term in the SM Higgs potential.

As we have seen, there are two complex Higgs doublets, H_u and H_d, present in the MSSM. When they acquire vacuum expectation values, this will cause spontaneous electroweak symmetry breaking. The scalar Higgs potential V_H in the MSSM comprises three components.

1. Quartic Higgs interactions resulting from the second term in equation 4.11. The $U(1)_Y$ part is denoted by V_D^1, the $SU(2)_L$ part by V_D^2.

$$V_D^1 = \frac{1}{2}\left[\frac{g_1}{2}\left(|H_u|^2 - |H_d|^2\right)\right]^2 , \tag{4.20}$$

$$V_D^2 = \frac{1}{2}\left[\frac{g_2}{2}\left(H_d^{i*}\tau_{ij}^a H_d^j + H_u^{i*}\tau_{ij}^a H_u^j\right)\right]^2 , \tag{4.21}$$

   with $\tau^a = 2T^a$ and $\tau_{ij}^a\tau_{kl}^a = 2\delta_{il}\delta_{jk} - \delta_{ij}\delta_{kl}$. Upon adding these two contributions, we find,

$$V_D = \frac{g_2^2}{8}\left[4|H_d^\dagger \cdot H_u|^2 - 2|H_d|^2|H_u|^2 + (|H_d|^2)^2 + (|H_u|^2)^2\right] + \frac{g_1^2}{8}\left[|H_u|^2 - |H_d|^2\right]^2 . \tag{4.22}$$

2. Contributions arising from derivatives of the superpotential, cf. Eqs. 4.9 and 4.10,

$$V_F = \sum_i \left|\frac{\partial \mathcal{W}(\phi_i)}{\partial \phi_i}\right|^2 = \mu^2\left(|H_d|^2 + |H_u|^2\right). \tag{4.23}$$

3. Soft SUSY-breaking scalar Higgs masses and bilinear terms, cf. Eq. 4.18,

$$V_{\text{soft}} = m_{H_u}^2 H_u^\dagger H_u + m_{H_d}^2 H_d^\dagger H_d + B\cdot\mu(H_u \cdot H_d + h.c.). \tag{4.24}$$

Expanding the Higgs fields into their components, we find for the total scalar

potential,

$$V_H = \left(|\mu|^2 + m_{H_u}^2\right)\left(|H_u^0|^2 + |H_u^+|^2\right) \tag{4.25}$$
$$+ \left(|\mu|^2 + m_{H_d}^2\right)\left(|H_d^0|^2 + |H_d^-|^2\right) \tag{4.26}$$
$$+ b\left(H_u^+ H_d^- - H_u^0 H_u^0 + h.c.\right) \tag{4.27}$$
$$+ \frac{1}{8}\left(g_1^2 + g_2^2\right)\left(|H_u^0|^2 + |H_u^+|^2 - |H_d^0|^2 - |H_d^-|^2\right)^2 \tag{4.28}$$
$$+ \frac{1}{2} g_2^2 |H_u^+ H_d^{0*} + H_u^0 H_d^{-*}|^2, \tag{4.29}$$

where we defined $b = B\mu$.

The minimum of this scalar potential should spontaneously break the electroweak symmetry down to electromagnetism $SU(2)_L \times U(1)_Y \to U(1)_{EM}$, as was the case in the SM. Using the freedom to make gauge transformations, we can choose, without loss of generality, $\langle 0|H_d^-|0\rangle = 0$ at the minimum of the potential. One can check that a minimum satisfying $\frac{\partial V_H}{\partial H_d^-} = 0$, must have $\langle 0|H_u^+|0\rangle = 0$. At the minimum, electromagnetism is thus necessarily unbroken, as the charged directions cannot attain a vacuum expectation value. Simplifying the scalar potential by setting $H_u^+ = H_d^- = 0$, we find,

$$V_H = \left(|\mu|^2 + m_{H_u}^2\right)|H_u^0|^2 + \left(|\mu|^2 + m_{H_d}^2\right)|H_d^0|^2 - b\left(H_u^0 H_u^0 + h.c.\right) + \frac{1}{8}\left(g_1^2 + g_2^2\right)\left(|H_u^0|^2 - |H_d^0|^2\right)^2. \tag{4.30}$$

Only the b term depends on the phases of the fields. This means that a possible phase in b can always be absorbed by a redefinition of the phases of H_u or H_d. Hence, we can take b to be real and positive. A consequence of this is that CP cannot be spontaneously broken by the Higgs scalar potential, at least at tree level. The Higgs scalar mass eigenstates will also have well-defined eigenvalues of CP.

To have a viable theory, we need to make sure that the potential is bounded from below, such that the vacuum is stable. In the Standard Model this was done by requiring the quartic Higgs coupling to be positive. Here, the quartic interactions in the potential will also stabilize the potential for large values of H_u^0 and H_d^0. Only in the special case $|H_u^0| = |H_d^0|$, corresponding to a flat direction in field space, do the quartic terms vanish. To ensure that the potential is still bounded from below, even in these cases, we find the requirement,

$$2b < 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2. \tag{4.31}$$

To have electroweak symmetry breaking, we need a linear combination of H_u^0 and H_d^0 to have a negative squared mass term. For this to happen the mass matrix should have a negative determinant,

$$\det\left(\frac{\partial^2 V_H}{\partial H_i^0 \partial H_j^0}\right) < 0. \tag{4.32}$$

This gives the requirement,

$$b^2 > \left(|\mu|^2 + m_{H_u}^2\right)\left(|\mu|^2 + m_{H_d}^2\right) \ . \tag{4.33}$$

If this inequality is not satisfied, $H_u^0 = H_d^0 = 0$ will be a stable minimum of the potential and there will be no electroweak symmetry breaking.

It is very interesting to note that both inequalities can only be satisfied at the same time when $m_{H_u} \neq m_{H_d}$. To break the electroweak symmetry we necessarily need to break supersymmetry. In some SUSY models both Higgs masses are assumed to be equal at some high unification scale, which means that there is no electroweak symmetry breaking. However, due to radiative corrections to the renormalization group equations, the masses at lower momentum scales will differ. These quantum corrections trigger electroweak symmetry breaking. This mechanism is often referred to as *radiative electroweak symmetry breaking*, and is viewed to be more natural than in the SM.

Assuming that the above stated conditions can be satisfied, we can now require the vev's to be compatible with the observed phenomenology of electroweak symmetry breaking. We write $v_u = \langle H_u^0 \rangle$ and $v_d = \langle H_d^0 \rangle$. They are related to the mass of the Z-boson

$$v_u^2 + v_d^2 = v^2 = \frac{2m_Z^2}{g_1^2 + g_2^2} \tag{4.34}$$

The ratio of the vev's is written as $\tan\beta = \frac{v_u}{v_d}$ and is not fixed by current experiments. The minimization equations, $\frac{\partial V_H}{\partial H_u^0} = 0$ and $\frac{\partial V_H}{\partial H_d^0} = 0$, can be written as

$$m_{H_u}^2 + |\mu|^2 = b\cot\beta + \frac{m_Z^2}{2}\cos(2\beta) \tag{4.35}$$

$$m_{H_d}^2 + |\mu|^2 = b\tan\beta - \frac{m_Z^2}{2}\cos(2\beta)\ . \tag{4.36}$$

Using these equations we can write m_Z and $\tan\beta$ as a function of b and $|\mu|$. Or the other way round, we can always eliminate b and $|\mu|$ and replace them with the variable $\tan\beta$. Only the sign of μ remains undetermined.

The Higgs scalar fields of the MSSM are two complex $SU(2)_L$ doublets, or eight real scalar degrees of freedom. Three of these are used to give mass to the W and Z bosons. After electroweak symmetry breaking five Higgs scalar mass eigenstates remain: two CP-even neutral scalars h^0 and H^0, one CP-odd neutral scalar A^0 and two charged scalars H^+ and H^-. By convention h^0 is lighter than H^0. To compute their masses, one can expand the doublet fields around their vacuum expectation value and plug this expansion into the Higgs potential. One finds that the mass of h^0 is bounded from above by $m_Z|\cos 2\beta|$ at tree level. Including radiative loop corrections and assuming that the sparticles contributing to the loop diagrams are lighter than about 1 TeV, this bound is $m_{h^0} < 135\,\text{GeV}$. This lightest Higgs boson is also expected to behave as the single Higgs from the SM for most of the MSSM parameter space. These predictions are thus consistent with the newly observed Higgs boson, and this strengthens the belief in SUSY as a possible SM extension. Using the observed Higgs mass, we can also invert this reasoning, and obtain limits on the masses of sparticles running in the loops. An example of this will be given in Section 4.6 when we discuss Natural SUSY. The masses of the other four Higgs bosons can, in principle, become arbitrarily large, and can thus escape detection at the LHC.

4.5 Sparticle production and decay

4.5.1 Mass eigenstates

In the MSSM, the superpartners listed in Tables 4.1 and 4.2 are not necessarily the mass eigenstates of the theory. After electroweak symmetry breaking, gauge eigenstates with the same $SU(3)_c \times U(1)_{EM}$ quantum numbers can mix with each other. The mass eigenstates can be determinant by diagonalizing the mass matrices, which can be constructed by gathering all terms in the Lagrangian that are quadratic in the fields.

The charged electroweak gauginos, $\widetilde{W}_1$ and $\widetilde{W}_2$ (the winos), and the charged higgsinos, $\widetilde{H}_u^+$ and$\widetilde{H}_d^-$, mix to form the so-called *charginos*, denoted by $\widetilde{\chi}_1^\pm$ and $\widetilde{\chi}_2^\pm$. The same happens for their neutral counterparts. The neutral wino $\widetilde{W}_3$, the bino $\widetilde{B}$, and the higgsinos, $\widetilde{H}_u^0$ and $\widetilde{H}_d^0$, mix to form *neutralinos*, denoted by $\widetilde{\chi}_{1,2,3,4}^0$. Neutralinos are very interesting for cosmology. If the lightest supersymmetric particle is a neutralino, this is an ideal candidate for dark matter: it is heavy, only weakly interacting and stable.

The gluino, which is a colour octet, cannot mix with any other particle. In this respect it is unique among all MSSM sparticles. The mass eigenstate is thus the same as the gauge eigenstate, and has as mass term M_3, coming from the soft SUSY breaking part of the Lagrangian (Eq. 4.16). The physical mass can be computed from the running of M_3 by considering the renormalization group equations.

The squarks and sleptons are scalars with the same quantum numbers, and so they can mix with each other. The mass eigenstates of the squarks and sleptons should therefore be determined by diagonalizing three 6×6 mass matrices – one for up-type squarks, one for down-type squarks and one for charged sleptons – and one 3×3 mass matrix – for the sneutrinos. However, because we usually assume that the soft breaking parameters are flavour-blind, the mixing angles are generally small. The exception are the third generation squarks and sleptons, which can have very different masses compared to their first and second generation counterparts, because of the effects of the large Yukawa and soft couplings for the third generation. Furthermore, they can also have substantial mixing in pairs $(\widetilde{t}_L, \widetilde{t}_R)$, $(\widetilde{b}_L, \widetilde{b}_R)$ and $(\widetilde{\tau}_L, \widetilde{\tau}_R)$. The off-diagonal terms in the mass matrices are given by,

$$m^2_{\widetilde{f}_{L,R}} = m_f \left(A_f + \mu \cot\beta\right), \tag{4.37}$$

for the up-type quarks and by,

$$m^2_{\widetilde{f}_{L,R}} = m_f \left(A_f + \mu \tan\beta\right), \tag{4.38}$$

for the down-type quarks and charged fermions. Because of the dependence on the fermion masses, the mixing is especially important for the top squark sector. The mass eigenstates are usually denoted by $\tilde{t}_1$ and $\tilde{t}_2$. The $\tilde{t}_1$ can become substantially lighter than all other squarks because of the large mixing that can be present. It is thus very well possible that the $\tilde{t}_1$ is the only accessible squark at the LHC. Mixing for the sbottoms and staus can become significant in the

case of large $\tan\beta$, while mixing for the first and second generation sfermions is insignificant due to the smallness of the electron, muon and light quark masses. The squarks of the first and second generation are usually grouped together, and assumed to be mass-degenerate.

4.5.2 Sparticle decays

The decay possibilities for SUSY particles depend strongly on the precise SUSY mass spectrum. This spectrum is very model-dependent, and is largely related to how SUSY has been broken, and which relationships between masses and couplings are assumed to be realized.

If kinematically allowed, two-body decays will always dominate. In case the mass spectrum is more compressed, and the direct two-body decay is forbidden, the decay will proceed via virtual sparticles giving rise to three- and four-body decays. These scenarios with small mass splittings can be very challenging to detect experimentally. In the following I will go through all sparticles and mention the decay options. The presence of antiparticles will be implicitly understood to reduce clutter in the notation.

Let us first consider the decays of the neutralinos and charginos. They contain an admixture of the electroweak gauginos and therefore inherit their couplings. Possible two-body decays are the decays to lepton + slepton, quark + squark and neutralino/chargino + $W/\mathrm{Z}/h$. If these two-body decays are not allowed, the three-body decay to fermion + fermion + neutralino/chargino will occur. In case the sleptons are relatively light, they will mediate the three-body decay, and the resulting fermions will often be leptons. In case the decay proceeds through the lightest Higgs boson, these fermions are often b quarks. This will then give rise to b tagged jets in the final state.

Sleptons decay predominantly to a lepton and a neutralino or chargino. In particular, the direct decay to lepton + LSP is hardly ever forbidden. The mass spectrum would have to be very compressed for this to happen, and is only really possible for staus if the mass difference between stau and LSP is less than the mass of the tau (1.77 GeV). If this is the case, the stau will be long-lived.

For squarks, the decay to quark + gluino will dominate if the gluino is light enough. The reason for this is the large QCD coupling. In case this decay is not accessible, the decay can proceed via the electroweak interaction to a quark + neutralino/chargino. For right-handed squarks the decay to the lightest neutralino will dominate if it is predominantly bino.

Finally, gluinos can only decay through a squark, which can be on-shell or not. As discussed above, the stop and sbottom can be much lighter than the other squarks, which means that $\widetilde{g} \to t\tilde{t}_1$ and $\widetilde{g} \to b\widetilde{b}_1$ could be the only available two-body decays. In this case, they will dominate and b-tagged jets will appear in the final states. In case all squarks are heavier than the gluino, it will decay via a three body decay to two quarks and a neutralino or chargino.

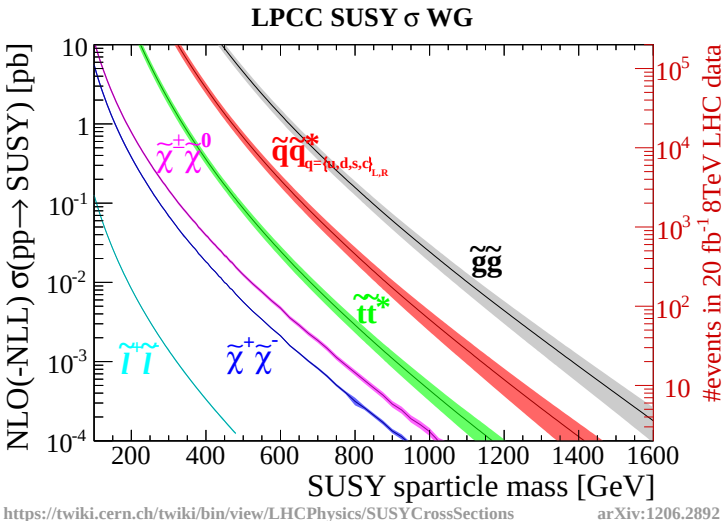

Figure 4.1: Cross sections for the production of various supersymmetric particles at a centre-of-mass energy of 8 TeV. The expected number of produced events for a dataset of 20 fb^{-1} are also shown. Figure taken from [43].

4.5.3 Signature at hadron colliders

When we assume that R-parity is conserved, the production of sparticles at hadron colliders such as the LHC can only occur in pairs. The production of squarks and gluinos will proceed predominantly via the strong interaction,

$$gg \rightarrow \tilde{g}\tilde{g},\ \tilde{q}_i\tilde{q}_j^*, \tag{4.39}$$
$$gq \rightarrow \tilde{g}\tilde{q}_i, \tag{4.40}$$
$$q\overline{q} \rightarrow \tilde{g}\tilde{g},\ \tilde{q}_i\tilde{q}_j^*, \tag{4.41}$$
$$qq \rightarrow \tilde{q}_i\tilde{q}_j, \tag{4.42}$$

while neutralinos, charginos and sleptons can only be produced via the electroweak interaction,

$$q\overline{q} \rightarrow \tilde{\chi}_i^+\tilde{\chi}_j^-,\ \tilde{\chi}_i^0\tilde{\chi}_j^0,\ \tilde{l}_i^+\tilde{l}_j^-,\ \tilde{\nu}_l\tilde{\nu}_l^*, \tag{4.43}$$
$$u\overline{d} \rightarrow \tilde{\chi}_i^+\tilde{\chi}_j^0,\ \tilde{l}_L^+\tilde{\nu}_l, \tag{4.44}$$
$$d\overline{u} \rightarrow \tilde{\chi}_i^-\tilde{\chi}_j^0,\ \tilde{l}_L^-\tilde{\nu}_l^*. \tag{4.45}$$

An overview of the cross sections for the production of various sparticles is shown in Fig. 4.1. The gluino-gluino production is seen to dominate, followed by the squark production. The electroweakinos have a much lower cross section. In principle, one can also have associated production of a chargino or neutralino together with a squark or gluino, the predicted cross-sections are usually much lower than for the other processes.

At the LHC, the production of squarks and gluinos is dominated the gluon-gluon and gluon-quark fusion. Once produced, they decay as explained in the

previous section, possibly resulting in a long decay cascade in which quarks and leptons can be produced. The full decay chain ends with the production of at least two LSP's due to the conservation of R-parity. When this LSP is a neutralino, it will escape from the detector without being detected. This will result in an imbalance of transverse momentum, which is one of the most powerful discriminating variables in the search for SUSY signals.

In general, SUSY signatures can contain any number of jets, leptons, and gauge bosons, depending on the details of the mass spectrum. The SUSY search program at the CMS and ATLAS experiments therefore covers a wide variety of final state topologies. Each analysis will be more or less sensitive to certain classes of SUSY parameter points, and it is thus of extreme importance to search everywhere, in all possible ways. The results of these searches can then be interpreted as limits on the allowed SUSY parameter space.

4.6 Natural SUSY

Let us now come back to the hierarchy problem. As we have discussed before, radiative corrections to the Higgs boson mass have the tendency to drive this mass up to very high scales in the SM. When supersymmetry is added to the theory, the quadratic contributions disappear due to the symmetries with the fermions and bosons that run in the loop. However, there are still logarithmic dependencies on the cutoff scale when SUSY is broken and the masses of the superpartners differ. The main contribution to the Higgs mass correction comes from the top quark and top squark loops, due to the large Yukawa coupling. One can show that the top squark contribution has the general form,

$$\Delta m_{H_u}^2 \propto -\frac{3}{8\pi^2} y_t^2 \left(m_{Q_3}^2 + m_{\mathrm{u}_3}^2 + |A_t|^2\right) \log \frac{\Lambda}{1\,\mathrm{TeV}}. \tag{4.46}$$

Requiring that this term does not become too large, results in a bound on the stop mass of roughly one TeV. A similar reasoning can be made for the gluino, which contributes at one-loop level to the top squark correction, and thus at two-loop level to the Higgs mass correction. The bound for gluinos to be still considered natural is around 1.5 TeV. The μ parameter also has a big effect on the Higgs boson mass corrections, as it contributes already at tree-level. The bound on μ, and thus the higgsinos, is of the order of 300 GeV.

With these considerations in mind, *natural SUSY* [44, 45] can be defined as the subset of the SUSY parameter space for which the above constraints on the sparticle masses are fulfilled. A natural SUSY scenario has low finetuning, and thus resolves the hierarchy problem. A typical mass spectrum is shown in Fig. 4.2. The idea of natural SUSY has driven many searches at the LHC, and has drawn particular attention to searches for gluinos and third generation squarks, as those particles are the only strongly produced particles that are required to be relatively light.

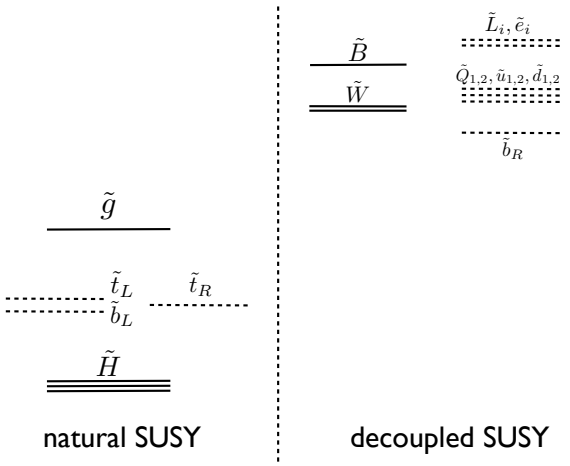

Figure 4.2: A natural SUSY spectrum, containing light higgsinos, light top squarks, and a light gluino. All other superpartners can be decoupled at high masses. Figure taken from Ref. [45].

4.7 Simplified model spectra

Many extensions of the Standard Model are possible, and many of the models have numerous free parameters. The MSSM is no exception. Without additional simplifications there are 104 extra free parameters; with a minimal set of assumptions, there are still about 20 parameters left. This extremely large parameter space is not very practical to do an analysis. Furthermore, similar experimental signatures can be produced in different ways, and by different models. To address these issues, a simplified approach was developed, using the so-called *simplified model spectra* (SMS) [46–48], rather than full models. These SMS are based primarily on the experimental topologies, how many jets or leptons are produced, whether there are jets originating from b quarks, etcetera.

Simplified models only contain a limited number of particles and interactions, and can be modelled as an effective theory. The result is a very minimal set of parameters for which no assumptions on relations between parameters have been made. This makes the simplified models both simpler, and more general. They are more suitable to optimize and interpret a new physics search compared to a full-fledged new physics model, for which, depending on assumptions, key topologies could be missing. The purpose of simplified models is three-fold,

- Identifying the boundaries of the search sensitivity. By scanning different masses, or mass differences between particles, it becomes easier to identify where searches lose sensitivity. This can then be used as input in the

proposal of new dedicated analyses.

- Characterizing possible new physics signals. If a signal is observed, SMS's can be used as a starting point to quantify the compatibility with different kinds of processes. Starting the characterization with full models would be very cumbersome due to the many free parameters that would need to be scanned.

- Derive limits on more general models. Complex models can be decomposed into experimental final state topologies. Ideally, each of those topologies would correspond to a SMS. The exclusion limits on the simplified models can then be translated back into a limit on the full model. Several efforts in this respect exist within the phenomenology community, such as SModelS [49, 50], Fastlim/ATOM [51], and checkMATE [52].

Let us consider as example a simplified model that only includes a gluino and the lightest neutralino. We further assume that the gluino can only decay as $\tilde{g} \rightarrow \mathrm{q}\overline{\mathrm{q}}\tilde{\chi}_1^0$. Simplified models are described by effective theories, in this example the decay proceeds through the dimension-six operator,

$$\mathcal{L}_{\mathrm{int}} = \frac{\lambda_i^2}{M_i^2}\tilde{g}\mathrm{q}_i\overline{\mathrm{q}}_i\tilde{\chi}_1^0 + h.c., \tag{4.47}$$

where i runs over the different quark flavours, λ_i is the Yukawa coupling for the quark-squark-$\tilde{\chi}_1^0$ vertex, and M_i is the effective scale of the interaction. Collision events can be simulated according to this effective Lagrangian in two main ways. The first is through the use of programs such as MARMOSET [53], which allows to directly simulate the on-shell effective theory. The second, and most widely used, way is to use a matrix element generator, such as MadGraph (see Section 6.2.1), to simulate events using the MSSM as model, setting the masses for all particles that are not involved in the SMS to very high values of $\mathcal{O}(100\,\mathrm{TeV})$. Any diagrams involving particles not included in the SMS will then effectively not contribute at all, and the MSSM will be reduced to the SMS under study. No matter the details of the simulation, the only parameters that are relevant for this SMS are the gluino production cross section, the branching ratio for the decay (here assumed to be 100%), and the masses for the gluino and the $\tilde{\chi}_1^0$. Analyses will then usually make a two-dimensional scan across the $(\tilde{g}, \tilde{\chi}_1^0)$ mass plane, and investigate how their sensitivity changes.

We also note that although the simplified models are SUSY-inspired, and use the SUSY nomenclature, they are in fact more general. Any model containing a spectrum of new narrow resonances, with the same gauge and flavour quantum numbers as the SM, and that are odd under some conserved parity, fits within the scope of the simplified models. Non-SUSY examples are the little Higgs models with T-parity, or universal extra dimensions with KK-parity.

Chapter 5

The Large Hadron Collider and the Compact Muon Solenoid experiment

The analysis presented in this thesis uses *pp* collision data delivered by the Large Hadron Collider (LHC) and recorded by the Compact Muon Solenoid (CMS) experiment. An overview of the collider is given in Section 5.1, while the detector setup is discussed in Section 5.2.

5.1 The Large Hadron Collider

The Large Hadron Collider is the world's largest particle accelerator and collider, able to reach unprecedented particle energies. It is located at CERN, the European Organization for Nuclear Research on the Swiss-French border near Geneva, in the 27 kilometre long tunnel that previously housed the Large Electron Positron (LEP) collider. The LHC consists of a sequence of superconducting magnets that guide two proton beams in opposite directions around the LHC ring, which is composed of 8 straight and 8 curved sections. The beams are accelerated at each turn around the ring, and are made to collide in four interaction points. If operating at design conditions, the LHC would provide 600 million *pp* collisions per second, at a centre-of-mass energy of 14 TeV and a luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. In the next sections I will highlight some of the main features of the LHC. A more comprehensive discussion can be found in Refs. [54–57].

5.1.1 A proton machine

Both LHC beams contain protons, unlike the previous large accelerators LEP and Tevatron which collided electrons on positrons, and protons on antiprotons, respectively. The decision to use solely protons was driven by the purpose of the LHC: to be a discovery machine. The LHC was built to test the Standard Model

at never seen energies, provide enough data to probe very rare processes, and hopefully discover new particles. So far the LHC has lived up to expectation. Per beam energies of up to 4 TeV were reached, providing access to a completely new energy domain. In 2012 the discovery of the long-sought-after Higgs boson was announced, followed in 2013 by a Nobel prize for François Englert and Peter Higgs, the theorists who first proposed the existence of this particle.

The advantage protons have over electrons and positrons is their 2000 times larger mass, resulting in a much reduced energy loss due to synchrotron radiation. This allows the LHC beam to reach energies that would be impossible at LEP. The drawback is the proton's complex structure. Electrons are fundamental particles, which means that the centre-of-mass energy of the collision is precisely known. This is not the case for protons. In a pp collision we know the energy of the protons, but not of the individual quarks and gluons inside the proton that participate in the hard interaction. Consequently, each collision has a different centre-of-mass energy, making it hard to perform precision measurements. The compositeness of the protons also leads to messier collisions, as the proton remnants can also interact and obscure the interesting hard collision event.

The LHC was built as a pp machine rather than a $p\overline{p}$ machine because antiprotons are hard to produce and unstable, thus limiting the maximal luminosity that can be achieved. Of course, there is also a downside here. A proton and antiproton beam can be circulated in opposite directions using the same magnetic field, and thus beam pipe. However, to steer two proton beams in opposite direction, we need to apply an opposite magnetic field, meaning that two separate beam pipes are needed.

5.1.2 LHC accelerator complex and experiments

Protons are not simply injected into the LHC itself, but rather pass through a series of pre-accelerators. It all starts with a bottle of hydrogen. Protons are obtained by applying an electromagnetic field to strip off the electrons from the hydrogen atoms. The protons pass through a linear accelerator, the Linac 2, where they obtain an energy of 50 MeV, and are then injected into the Proton Synchrotron (PS) Booster. The booster accelerates the protons to 1.4 GeV at which point the beam is fed to the PS where it attains an energy of 25 GeV. From the PS the beam is sent to the Super Proton Synchrotron (SPS) where the protons reach an energy of 450 GeV. The beams are then finally transferred to the LHC, both in a clockwise and an anticlockwise direction, where they are accelerated to their final energy. This entire process is illustrated in Fig. 5.1.

In addition to accelerating protons, the accelerator complex can also accelerate lead ions, which are produced from a highly purified lead sample, heated to a temperature of about 500°C. An electric current is used to ionize the lead vapour. The lead ions pass through Linac 3, and are then accumulated and accelerated to 72 MeV per nucleon in the Low Energy Ion Ring (LEIR). From the LEIR the ions are transferred to the PS, from where they follow the same path as the protons, reaching a final energy of 2.76 TeV per nucleon in the LHC.

Once inserted, beams will circulate inside the LHC beam pipes for many

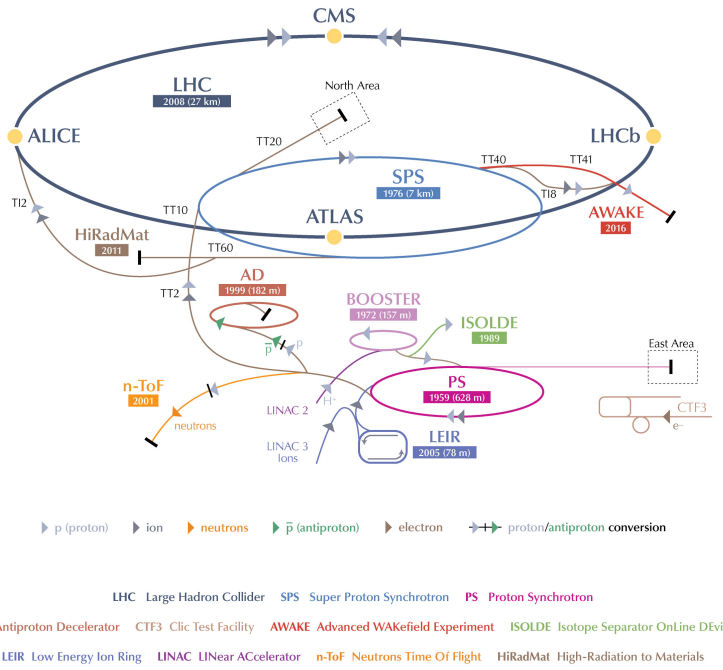

Figure 5.1: The CERN accelerator complex. Protons leave the Linac 2, are passed to the Booster, the PS and SPS, before being injected in the LHC. The four main LHC experiments, CMS, ATLAS, LHCb, and ALICE, are also shown. Figure taken from Ref. [58].

hours under normal operating conditions. During this time the two beams are made to collide in four interaction points, each containing a separate detector:

- **CMS** is a general purpose detector with a very broad physics programme covering Standard Model measurements, Higgs physics, searches for possible new particles, etcetera. Its main feature is the huge solenoid magnet, operating at a magnetic field strength of 3.8 T. The CMS detector will be explained in more detail in the next chapter.

- **ATLAS** is a general purpose detector like CMS, with a very similar physics programme, but utilizing different detector techniques and a magnet design featuring a toroid. At 46 m long, 25 m high and 25 m wide, the ATLAS detector is the largest volume particle detector ever constructed, although not as heavy as the CMS detector.

- **LHCb** is a more specialized experiment, aiming to unravel the mysteries of antimatter by studying processes involving b quarks. Unlike the cylindrical CMS and ATLAS detectors, the LHCb detector is asymmetric and targets detection of forward particles in particular.

- **ALICE** is a heavy-ion detector with as primary goal the study of the quark-gluon plasma, a state of matter occurring at extreme densities. It is the main customer of the LHC heavy-ion (lead) collision runs.

As the two beams collide with each other, or with the remaining gas in the ultrahigh vacuum of the beam pipe, the beam intensity keeps dropping. At a certain point more collisions can be delivered to the experiments by starting the full chain all over again. The beam will then be dumped, and a fresh set of proton bunches will be inserted. If at any time a magnet would *quench*, have an increase in temperature resulting in a loss of the superconducting state, the beams need to be dumped as well for safety reasons.

The beam dump system [60] is shown schematically in Fig. 5.2. Once the decision to dump the beam is made, a fast kicker magnet is turned on to deflect the beam in the horizontal plane. This kicker magnet has a pulse rise-time of 3 μs. To accommodate this rise-time, and allow for the safe extraction of the beam, the LHC beam structure is designed to have a gap, where there are no filled bunches. The kicker magnets guide the beam onto so-called septum magnets which fully extract the beam from the beam pipe, deflecting it vertically, and steer it toward the graphite beam dump blocks. The beam passes through a set of dilutor kicker magnets which spread the beam in both horizontal and vertical direction, tracing out an 'e' shaped path. The beam size increases from 0.2 mm to 1.5 mm upon reaching the dump blocks 700 m away. Without this dilution, the local beam intensity and heat production would be too large for the blocks to handle, they would vaporize immediately. However, the graphite core of the dump block still needs to endure temperatures of up to 750°C. The core has a cylindrical shape, 0.7 m in diameter and 7.7 m long, and is surrounded by about 900 tons of radiation shielding blocks.

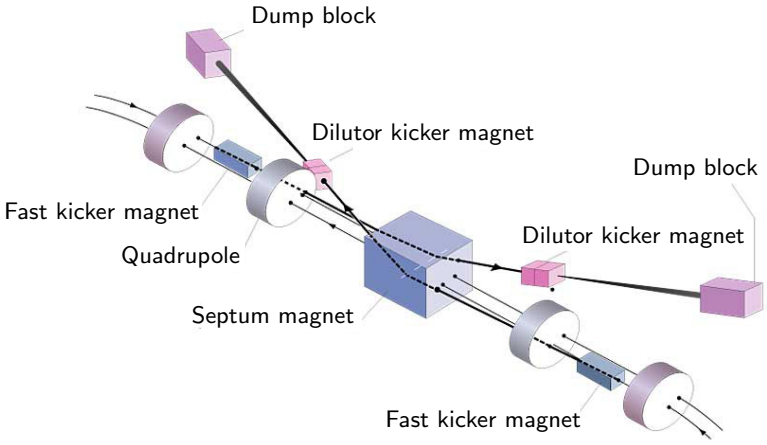

Figure 5.2: The LHC beam dump system. Figure adapted from Ref. [59].

5.1.3 Superconducting magnets

In order to reach the LHC design centre-of-mass energy of 14 TeV, conventional magnets do not suffice. Therefore, superconducting magnets are used instead. They can provide the very high magnetic fields, up to 8.4 T, needed to bend the highly energetic particles around the ring, at reasonable power consumption. The magnet coils are built from niobium-titanium cable, that when cooled to 1.9 K attains a superconducting state, allowing electricity to flow without resistance. The cryostat at the LHC is of unseen proportion, stretching along the 27 km long ring, and containing 120 tonnes of helium.

The need for two separate beam pipes caused some complication for the LHC design. The LEP tunnel is quite narrow, too narrow, in fact, to hold two completely separate proton rings. A solution was found in the so-called twin-bore, or two-in-one, magnet design. This design accommodates the windings for both beam channels in a common cold mass and cryostat, with the magnetic flux circulating in the opposite sense through the two channels. It also makes the magnet structure more complicated, as the separation between the two channels is small enough that they are coupled both magnetically and mechanically. An illustration of the twin-bore design is shown in Fig. 5.3.

Guiding beams around the LHC ring requires thousands of magnets of different varieties and sizes. Beams are bent by a total of 1232 dipole magnets, each 15 metres in length, and are focussed by 392 quadrupole magnets, each 57 metres long. Different kinds of multipole magnets are used to correct small imperfections in the magnetic field at the ends of the dipoles. So-called insertion magnets are used to squeeze the beam from 0.2 millimetres down to 16 micrometres across, just before it reaches one of the detectors in the interaction points.

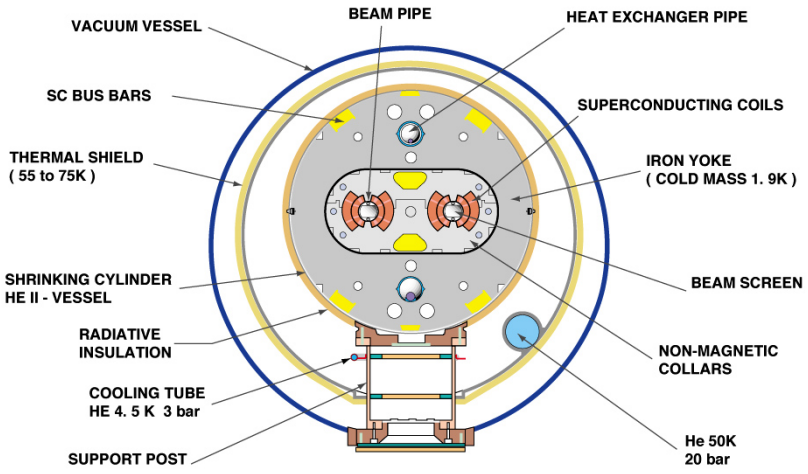

Figure 5.3: Cross section of an LHC dipole magnet, illustrating the twin-bore magnet design [61]. The two beam pipes surrounded by the dipole magnets are clearly visible. They are encased in a single cold volume and vacuum vessel.

5.1.4 Accelerating cavities

The necessary accelerating power to raise the beam energy from the injection energy of 450 GeV to the collision energy of several TeV is provided by eight radiofrequency (RF) cavities per beam.

RF cavities are metal chambers, often structured like beads on a string, where the beads are the cavities and the string is the beam pipe of the accelerator. An electromagnetic (EM) field is supplied to the cavities by an RF power generator. The specific shape and size of the cavities are such that the EM waves become resonant, and build up inside the cavity. Charged particles passing through the cavity feel the force and direction of the resulting electromagnetic field, and are pulled along with the field. As the field in the LHC RF cavities oscillates at 400 MHz, the arrival of the protons needs to be timed precisely in order for them to be accelerated, and not decelerated, see Fig. 5.4. Once the particles have circled the LHC ring a sufficient number of times, passing through the RF cavities on each turn, they attain the desired energy. At this point the main purpose of the RF cavities is to keep the protons inside their *bunches*.

The protons in the LHC beams do not form a continuous flow, but are grouped in packets, the bunches, with empty space in between. Up to 2808 bunches can circle the LHC at any given time. An ideally timed proton, with exactly the right energy, will see zero accelerating voltage from the RF cavities when the LHC is at full energy. Protons with slightly different energies arriving earlier or later will be accelerated or decelerated so that they stay close to the energy of the ideal particle. In this way, the beam bunches stay intact.

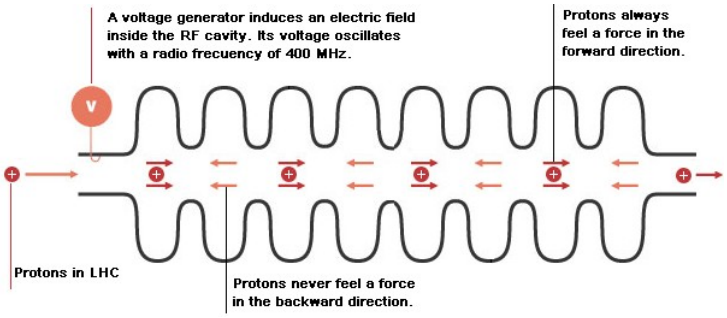

Figure 5.4: Interaction between a proton and the oscillating electromagnetic field inside a radiofrequency cavity. Figure taken from Ref. [62].

The LHC utilizes superconducting RF cavities made of niobium sputtered on copper, which are cooled to 4.5 K. The RF power, up to 2 MV per cavity, is delivered by klystrons and passed to the cavities via waveguides. Protons can thus gain up to 16 MeV per turn around the ring. The RF cavities are grouped by fours in cryomodules, which are installed in a straight section of the LHC tunnel.

5.1.5 Past and future running periods

The first proton beams circled the LHC in September 2008. Unfortunately, an incident happened shortly after, during powering tests of the dipole magnet. A faulty connection caused an electrical arc which punctured the helium enclosure, leading to a release of helium into the insulation vacuum of the cryostat. The pressure from the expanding helium rose too fast for the relief valves to cope, thereby damaging dozens of magnets. It took more than one year to repair all the damage, and install extra safety systems.

In November 2009 beams were back in the LHC, with first stable beams at 7 TeV centre-of-mass energy in March 2010. The LHC continued to run very smoothly throughout 2011, delivering data corresponding to an integrated luminosity of $5\,\text{fb}^{-1}$ to the experiments. In 2012 the beam energy was increased from 3.5 to 4 TeV, and also the luminosity of the beam saw a sharp increase. A total integrated luminosity of over $20\,\text{fb}^{-1}$ was delivered.

During 2013 and 2014 the LHC was shut down for maintenance and upgrade, in order to prepare the machine to run at 13 TeV centre-of-mass energy and higher instantaneous luminosity. During these two years, the experiments at the interaction points were also upgraded to increase detector performance in light of the changing conditions expected upon restart. First stable beams for physics collisions at 13 TeV are planned for June 2015. This will signal the start of a very exciting time for particle physics, during which our understanding of elementary particles and their interactions will undoubtedly change.

5.2 The Compact Muon Solenoid experiment

The Compact Muon Solenoid (CMS) [63–66] is one of the two general purpose detectors at the LHC. It is located on the far-end of the LHC ring, at interaction point 5 (P5) in Cessy, France. Like most collider experiments, CMS has a cylindrical shape, consisting of a barrel region and two so-called endcaps at either end of the barrel. The various subdetector systems are layered around the collision point. The central feature of the CMS detector is the superconducting solenoid with high magnetic field to achieve good momentum resolution. The muon systems are installed in between the return yoke layers. The silicon tracker, lead-tungstate crystal electromagnetic calorimeter (ECAL), and brass scintillator hadron calorimeter (HCAL) are contained within the bore of the magnet coil. The CMS detector has a length of 21 m, a diameter of 15 m and a total weight of 14 000 tonnes. Figure 5.5 shows an overview of the detector and its different subsystems. It is important to note that different kinds of particles interact differently with the various subdetectors, as illustrated in Fig. 5.6. This design allows the reconstruction software to distinguish between electrons, muons, photons, charged and neutral hadrons. A more extensive discussion on each subsystem will be given in the following sections.

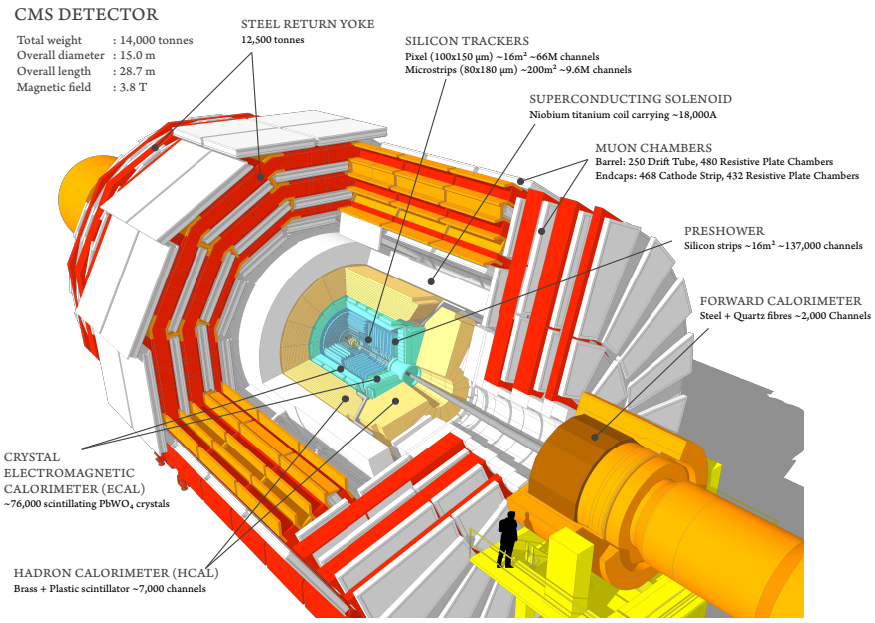

Figure 5.5: Overview of the CMS detector. The different subsystems are indicated on the figure, as well as a person to serve as reference scale. Figure taken from Ref. [67].

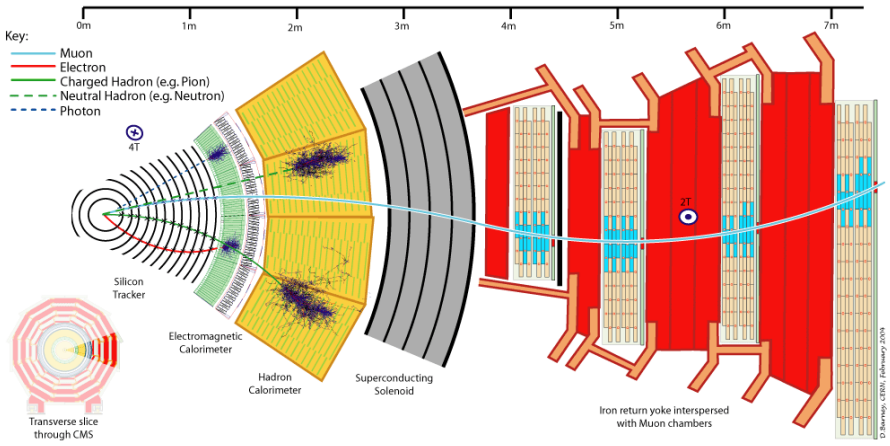

Figure 5.6: Detailed view of a CMS detector slice, showing the interactions of different kinds of particles with the various subsystems. Muons leave hits in the tracker and the muon stations, before leaving the detector. Electrons leave hits in the tracker, and then deposit their energy in the ECAL. Photons can be identified as an energy deposit in the ECAL without a corresponding track. Charged and neutral hadrons both deposit their energy in the HCAL, with matching tracker hits for charged hadrons only. Figure taken from Ref. [68].

5.2.1 Coordinate system and basic variables

The CMS coordinate system takes the nominal collision point as the origin, with the y-axis pointing vertically upward, and the x-axis pointing radially inward towards the centre of the LHC. The coordinate system is right-handed, such that the z-axis points along the beam direction toward the Jura mountains as seen from P5. The azimuthal angle ϕ is measured from the x-axis in the (x, y) plane. The polar angle θ is measured from the z-axis. Pseudorapidity η is defined as

$$\eta = -\ln\left[\tan(\theta/2)\right], \tag{5.1}$$

and is usually used instead of the polar angle because of the property that differences in pseudorapidity are Lorentz-invariant. This is especially useful for hadron colliders where the boost in the z direction is unknown, and varies for each collision. A Lorentz-invariant angular separation between two particles is defined as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{5.2}$$

The momentum and energy measured transverse to the beam direction, denoted by p_{T} and E_{T}, respectively, are computed from the x and y components. The imbalance of energy measured in the transverse plane is denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$, and indicates undetected particles, or energy mismeasurements.

5.2.2 Magnet system

The CMS magnet is a superconducting solenoid, producing a uniform magnetic field of 3.8 T inside the magnet coil. Having a magnet is key to the success of any collider experiment. Without the bending of charged-particle tracks in the magnetic field, it would be very hard, if not impossible, to obtain accurate momentum and charge measurements for those particles. Since the curvature decreases as the p_{T} of the particles increases, the strength of the magnetic field, and the precision of the tracker, will determine how accurately the momentum of highly energetic particles can be measured.

To take as much advantage of track bending as possible, CMS decided to have the strongest magnetic field possible. Therefore, the largest magnet that could be transported to P5 was built. The CMS solenoid measures 6.3 m in diameter, and the steel return yoke has an outer diameter of 14 m. The yoke consists of three layers, as shown in Fig. 5.7, provides most of the structural support for the whole experiment, and acts as a shield that blocks any particles that made their way across the HCAL, apart from muons, neutrinos, and possible new weakly interacting particles. The total magnet system is 13 m long, weighs a whopping 12 000 tonnes, and is the largest superconducting magnet ever built. The tracker, ECAL, and HCAL fit inside the magnet coil, whereas the muon stations are interleaved with the return yoke.

The solenoid is cooled to 4.5 K, and the very strong magnetic field of 3.8 T results in a momentum resolution of $\Delta p/p \approx 10\%$ at momenta of 1 TeV, sufficient to determine the sign of muons with $p_{\mathrm{T}} \approx 1$ TeV.

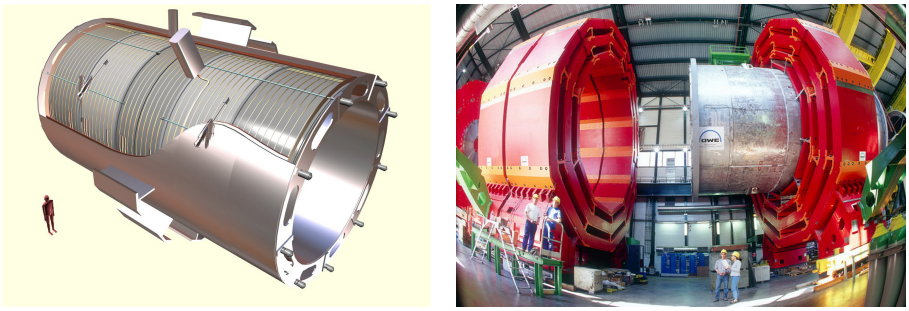

Figure 5.7: [left] Artistic view of the superconducting solenoid showing the five modules composing the cold mass inside the cryostat. Figure taken from Ref. [63]. [right] Fish-eye view of the red magnet yoke during construction in 2002. Figure taken from Ref. [69].

5.2.3 Tracker

The purpose of the tracker is to very accurately measure the curved tracks of charged particles with $p_{\mathrm{T}} > 1$ GeV, such that a precise momentum measurement can be made. The tracker should also be able to precisely reconstruct secondary vertices stemming from e.g. b quark decays. In addition, particles should be disturbed the least amount possible while passing through the

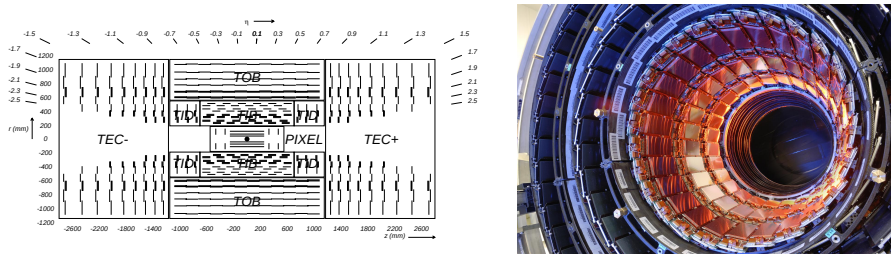

Figure 5.8: [left] Schematic cross section through the CMS tracker. Each line represents a detector module. Figure taken from Ref. [63]. [right] The first half of the CMS inner tracker barrel (TIB), consisting of three layers of silicon modules. Figure taken from Ref. [70].

tracker. The tracker should thus consist of as little material as possible in order to avoid multiple scattering, bremsstrahlung, photon conversion, and nuclear interactions. To accomplish this, the tracker comprises several thin layers. When a particle passes through a layer, it creates a signal, a *hit*. The reconstruction software then builds a track out of the different hits.

The tracker is the innermost layer of the CMS detector, perfectly placed to measure the particles coming directly from the collision point. Its closeness to the beam pipe also means that it will receive the largest amount of particles, and thus radiation, which translates into the need for radiation hard materials.

The CMS tracker is entirely built out of silicon, and consists of a pixel detector at the very centre, surrounded by a microstrip detector. It has a length of 5.8 m and a diameter of 2.5 m. With about 200 m^2 of active silicon area, the CMS tracker is the largest silicon tracker ever built. A diagram of the structure is shown in Fig. 5.8, alongside a photograph of part of the strip detector. There are a total of 75 million read-out channels with very fast response, providing a precision of 10 μm in position measurement, even when up to 1000 particles traverse the tracker every 25 ns. This corresponds to a transverse momentum resolution of 1-2% for $p_{\mathrm{T}} \approx 100\,\mathrm{GeV}$.

The pixel detector has three cylindrical layers very close to the collision point (at 4 cm, 7 cm and 11 cm from the beampipe), and two disks at either end. Each of the 65 million pixel sensors measures 100 μm by 150 μm. A cooling system is installed to keep the temperature from rising too much above the operating temperature of $-10°\mathrm{C}$.

The silicon strip detector consists of ten layers in total. There are four inner barrel (TIB) layers with two inner endcaps (TID), each composed of three small discs. The outer barrel (TOB) consists of six layers, while two endcaps (TEC) close off the tracker. The strip tracker is cooled to a temperature of $-20°\mathrm{C}$ in order to minimize the spreading of any radiation damage.

5.2.4 Electromagnetic calorimeter

The CMS electromagnetic calorimeter is a hermetic homogeneous calorimeter, consisting of a central barrel region and two endcaps, and is located in between

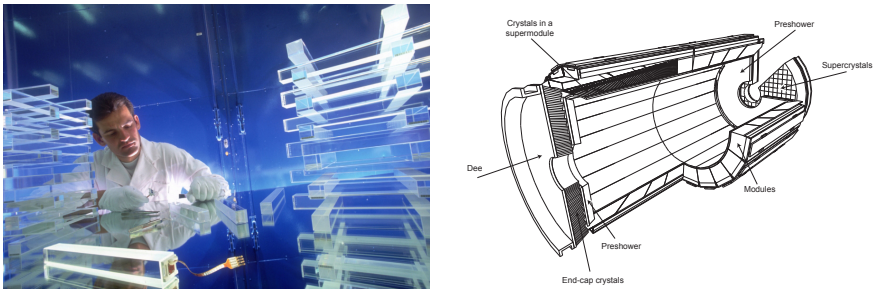

Figure 5.9: [left] The construction of the CMS electromagnetic calorimeter: lead-tungstate crystals being tested in 2005. Figure taken from Ref. [71]. [right] Schematic overview of the ECAL, showing the barrel modules that contain the crystals, the endcaps, and the preshower detectors. Figure taken from Ref. [63].

the tracker and the HCAL. It has a fine granularity, is fast, and radiation resistant. The ECAL provides a very good energy resolution for electrons and photons, which played a key role in the discovery of the Higgs boson in the $h \to \gamma\gamma$ final state.

The electromagnetic calorimeter is made of about 75 000 highly transparent lead tungstate ($PbWO_4$) crystals whose lengths corresponds to about 25 radiation lengths. A photograph of the crystals during testing, and a schematic view of the ECAL structure are shown in Fig. 5.9. The crystals are arranged side-by-side with their longitudinal axes slightly pointing away from the interaction point to avoid cracks.

When electrons or photons pass through such a lead tungstate crystal, it scintillates. This scintillation light is detected, and converted to an electrical signal, by avalanche photodiodes (APD) in the barrel, and vacuum phototriodes (VPT) in the endcaps. The scintillation decay time is very short, about 80% of the light is emitted within 25 ns, which is ideal for the bunch spacing of the LHC. The light yield of the crystals depends strongly on temperature, however. In order to keep the temperature at a stable level, to within 0.1°C, a dedicated temperature control system was put in place. Even though the crystals are quite radiation hard, they still suffer damage in the very high particle flux environment within CMS. Fortunately, the damage repairs itself through annealing during the periods when the accelerator is not operational, and the temperature of the crystals is higher. A light monitoring system is used to accurately measure the transparency of the crystals throughout the running period, so that precise energy measurements can always be made.

Preshower detectors in front of the ECAL endcaps provide for extra spatial precision, making it possible to distinguish between single high-energy photons and the less interesting close pairs of low-energy photons coming from the decay of neutral pions. The preshower detector is a sampling calorimeter, consisting of two planes of lead radiators that initiate showers when an electron of photon passes through. Behind each lead plane are sensors that measure the deposited energy and transverse shower profiles. The sensors are silicon strip detectors

of only 2 mm wide, compared to 3 cm for the ECAL endcap crystals. The two layers of strips are placed orthogonally, thus providing a position measurement with extremely fine granularity.

5.2.5 Hadron calorimeter

The hadron calorimeter is the last subdetector that is (mostly) located within the magnet coil. It is designed to detect and absorb hadrons, such that the only particles leaving the HCAL are muons and very weakly interacting particles such as neutrinos. For this purpose it was built to be as hermetic as possible, staggering the detector layers to make sure there are no gaps in straight lines that would allow hadrons to escape undetected. Without this hermeticity it would be impossible to use the missing transverse energy as a way to infer the presence of very weakly interacting particles. Many new physics searches, e.g. SUSY searches, would lose most of their sensitivity.

The HCAL consists of four subparts: the barrel (HB), outer barrel (HO), endcap (HE) and forward (HF) sections. The barrel and endcap subsystems are sampling calorimeters, made of alternating layers of dense absorber material (brass or steel) and fluorescent plastic scintillator tiles. When a hadronic particle hits an absorber plate, an interaction can occur producing numerous secondary particles. These secondary particles then flow through the successive layers of absorber material, where they too can interact, resulting in a cascade or shower of particles. As this shower develops, the particles pass through the layers of scintillators, causing them to emit blue-violet light. This light is absorbed by wavelength-shifting optical fibres, which shift the light into the green region of the spectrum, before being transported to the readout boxes by clear optical cables. The amount of light that is collected is a measure of the energy of the passing particle. Hybrid photodiodes convert the optical signals into fast electronic signals which are then sent to the data acquisition system. A photograph of the barrel HCAL is shown on the left in Fig. 5.10.

To fully absorb a particle shower, about one metre of absorber material is needed. As CMS is a very compact detector, it was impossible to fit the full HCAL barrel inside the magnet coil. This is the reason why the outer barrel, the tail-catcher, is located just behind the magnet coil, where it ensures that the punch-through particles are absorbed as much as possible, rather than create fake hits in the muon system.

The two hadronic forward calorimeters are positioned at either end of CMS, at 11.2 m from the interaction point, extending the coverage down to very high $|\eta|$. The HF receives the bulk of the particle energy contained in the collision, and so must be very resistant to radiation. The design of the HF was primarily guided by the necessity to survive the hostile conditions. Over 1000 km of quartz fibres are inserted into steel absorber blocks, as can be seen on the right in Fig. 5.10. When particles traverse these fibres, they generate Cherenkov light. This light is then transported through the fibres to photomultiplier tubes where it is converted into an electrical signal. The HF is divided into two longitudinal segments, making it possible to distinguish the more shallow showers generated by electrons and photons, from those generated by hadrons.

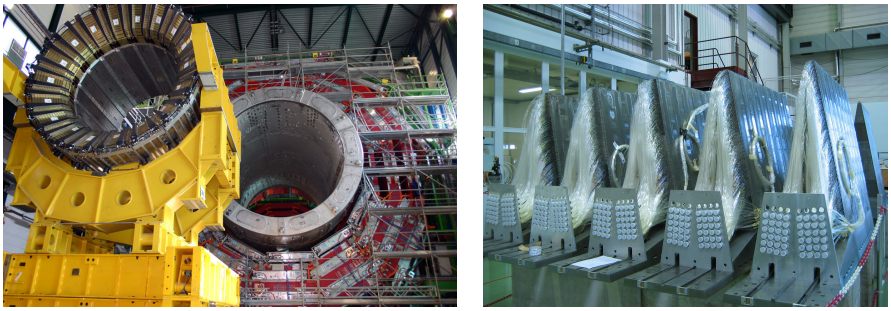

Figure 5.10: [left] Insertion of part of the barrel HCAL into the solenoid for the Magnet Test and Cosmic Challenge in 2006. Figure taken from Ref. [72]. [right] Segments of the forward HCAL during assembly, showing the quartz fibres inside the steel wedges. Figure taken from Ref. [73].

5.2.6 Muon system

Detecting muons, the heavier sibling of electrons, is one of the priorities of the Compact *Muon* Solenoid. Because it is 200 times heavier, a muon interacts less with material than an electron, enabling it to penetrate several metres of iron, and pass through the calorimeter systems unstopped. The muon chambers are, therefore, placed behind all the other subdetectors where muons are the only particles that would register a signal. This also means that muons are not easily mistaken for any other particle, thus providing a very clean signature. The Higgs boson decay to four muons is often referred to as the *golden channel* for exactly this reason, and it was indeed key to the Higgs boson discovery and the very accurate measurement of its mass.

The muon system consists of four muon stations, interleaved with the magnet return yoke, that track the path of a muon. The momentum of the muon is then inferred from its curvature in the magnetic field. Three different types of muon chambers are used: resistive plate chambers (RPC) provide very fast trigger information with rough positional information, while drift tubes (DT) and cathode strip chambers (CSC) provide accurate positional information and additional trigger information. In the barrel region, DTs and RPCs are arranged in concentric cylinders, whereas the endcap disks contain different layers of CSCs and RPCs. Figure 5.11 shows a schematic view of the muon system, along with a picture from the barrel muon chambers in the return yoke.

Drift tubes contain a stretched wire within a gas volume. Any charged particles passing through the gas create free electrons, which then drift towards the positively charged wire where they are registered. Each CMS DT chamber, on average 2 m by 2.5 m in size, consists of 12 aluminium layers, arranged in three groups of four, the superlayers, each containing up to sixty 4 cm wide drift tubes. The DTs can provide two coordinates for the position of the muon because the superlayers are placed orthogonally to each other. The (r, ϕ) coordinate is measured by wires parallel to the beampipe, while the z coordinate is obtained from wires perpendicular to the beampipe. How far from the wire

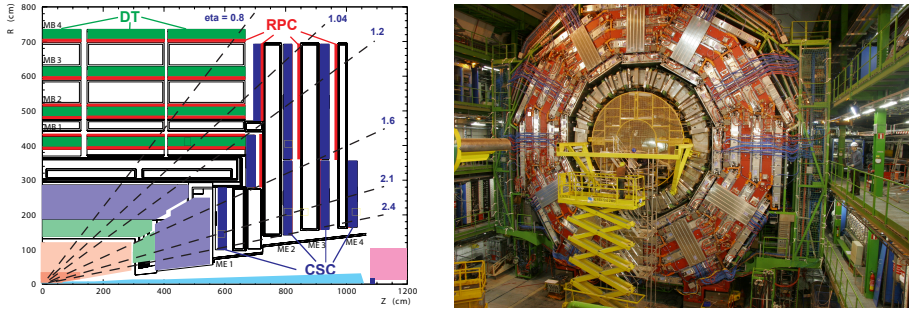

Figure 5.11: [left] Schematic overview of one quarter of the CMS muon system, showing the DT, CSC and RPC detectors. Figure taken from Ref. [64]. [right] Muon chambers interspersed in the magnet return yoke. Figure taken from Ref. [74].

the muon actually passed, is determined from the electron drift speed and the arrival time of the electrons at the wire.

Cathode strip chambers are more radiation resistant compared to drift tubes, and are, therefore, used in the endcap disks where the magnetic field is uneven and particle rates are high. The CSCs are so-called multiwire proportional chambers, consisting of positively-charged anode wires crossed with negatively-charged copper cathode strips inside a gas volume. Muons passing through one of the 468 trapezoidally shaped CSCs ionize the gas, creating electrons that move towards the anode wires, and positive ions that move towards the copper cathode. Because the strips and the wires are perpendicular, we obtain two position coordinates. The wires provide the radial coordinate r, and the strips the ϕ coordinate. Given that the wires are very closely spaced, the CSCs are fast and can thus be used as input to the trigger decision, besides performing the task of precision muon measurement.

Resistive plate chambers are made of two parallel plates of a very high resistivity material, a positively-charged anode and a negatively-charged cathode, which are separated by a gas volume, the *gap*. The CMS RPC system uses 610 double-gap modules comprising two gas gaps with readout strips in between. As for the other muon detectors, passing muons knock electrons out of gas atoms. These electrons then initiate an avalanche which moves towards the anode. The avalanche is detected by the metallic readout strips. The pattern of the hit strips gives a measure of the muon momentum. Because RPCs have a time resolution of just one nanosecond, they are very well suited be part of the muon trigger system.

5.2.7 Trigger and data acquisition

At design specifications, CMS detects a proton bunch crossing every 25 ns. This means that there are 40 million collisions per second, too much to fully read out, process and store for further analysis. Reducing this amount of data to a more manageable few hundred events per second is the role of the trigger

system. It is of course of prime importance to keep the interactions that could reveal new phenomena, such as supersymmetry, and discard those that provide little new information. The decision to keep or throw out an event is made in two stages, at the Level-One (L1) Trigger and the High-Level Trigger (HLT). Both of those systems will be discussed in the next sections.

The enormous data production rate has another consequence. The collisions occur so fast that the particles produced in one event have not yet left the detector when the next collision occurs. Associating each particle, and thus each detector signal, to the correct bunch crossing is only possible because of the very good time resolution of the detectors, and the synchronization between all the readout channels. Data from each subdetector is first collected in separate pipelines and is then sent off to the switch networks to build the full event. Section 6.4 will discuss how the events are then further reconstructed.

L1 Trigger

The L1 trigger is designed to make extremely fast decisions, based on relatively simple criteria, reducing the data rate to about 100 kHz.

When particles from a collision pass through the different subdetectors, they generate signals. These signals are collected in buffers in the front-end electronics of the subdetectors themselves. A subset of the information is immediately passed along to the L1 trigger, which is housed in a service cavern next to the detector. A total time of 3.2 μs is allocated for the transit to the L1 trigger boards, choosing whether to keep or reject the event, and the transit back to the front-end electronics to execute what was decided. During this time, the total amount of information on a given collision needs to be kept in the buffers. Less than 1 μs is allocated to the L1 trigger to perform the calculations needed to make the decision. It is, therefore, built from custom hardware processors, fully optimized for the task at hand.

Because decisions have to be made so quickly, it is impossible to use all available information, so only the calorimeters and the muon system are included. The trigger objects, such as photons, electrons, muons, and jets, are constructed by the global calorimeter trigger and global muon trigger, using the reduced granularity and resolution data sent to them from the front-end electronics. The objects are then passed to the Global Trigger, which makes the decision to keep an event or not based on the number of objects above given E_{T} or p_{T} thresholds, and/or on the total summed E_{T} or $E_{\mathrm{T}}^{\mathrm{miss}}$ of the event. If the event is to be kept, the high-resolution data is sent from the readout pipelines to the event builder and the High Level trigger.

High Level Trigger

The High Level trigger [75] is a computer farm that has access to the full event information, and can therefore perform complex calculations similar to those made in the offline analysis software. Each event reaching the HLT is sent to a separate processor which spends on average less than 0.1 s to make the final trigger decision, although for some events this can reach up to 1 s. The trigger software is set up in such a way that on average only a few hundred

events per second are kept and written to disk. A trigger path, a combination of object requirements that specifies which events should be kept, can be as complex as needed, as long as the computation is fast enough. To facilitate this, the trigger paths are set up in such a way that events are discarded as soon as possible, before reaching the more time consuming parts of the code. An example of this could be the use of jets constructed from only calorimeter information, before adding the information from the numerous tracker readout channels to improve precision. Each event that is kept receives a set of tags associated with the separate trigger paths that were fulfilled. These tags are later used to sort the data into so-called primary datasets, grouping together tags featuring similar objects.

Data acquisition

The CMS data acquisition (DAQ) system [76] comprises different components, which are illustrated in Fig. 5.12. The about 700 detector front-end drivers store the data from the detector front-end electronics upon the reception of an accept signal from the L1 Trigger. The data are then read by the readout system and stored until they are sent to the HLT for further processing. The builder network provides the interconnections between the readout and the HLT systems, running with a bandwidth of 100 GB/s. It takes care of building the full events out of the different data fragments from the separate parts of the detector. The builder network consists of 500 builder units, each of which will assemble the data for one event. Once the event is assembled, it is passed to one of the HLT processors associated to that particular builder unit. As soon as the HLT decision is made, that decision is transferred back to the builder unit which then either discards the event, freeing up memory, or sends it to the storage manager to be written to disk. Apart from building the events, the DAQ also provides detector control and monitoring services.

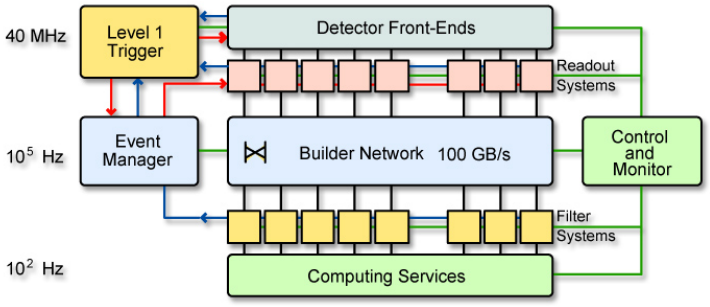

Figure 5.12: General architecture of the CMS DAQ System. Figure taken from Ref. [64].

Chapter 6

Event generation, simulation and reconstruction

In the previous chapter I discussed how *pp* collisions are produced by the LHC, and how they are detected by CMS. This chapter will elaborate on the different steps needed to actually use the collisions for physics analysis. First, I will explain more details about the collisions, or *events*, themselves. I will discuss the separate components an event consists of, as well as touch upon how events are described mathematically.

In order to understand what we observe in the data, which might contain signals of new physics, it is important to know how the Standard Model processes and new physics signals will appear in the detector. To achieve this we generate those processes using Monte Carlo generation techniques, incorporating everything that is known about the workings of the Standard Model, and the considered models of new physics. The principal options that are available to generate events will be discussed in Section 6.2. For each generated collision, we obtain a set of final state particles according to the specified physics process. This could for example be the particles that result from the production and decay of a top quark. At this stage we do not know yet how these particles would interact with the detector. That is taken care of in a next step by the event simulation, as explained in Section 6.3. An event simulator mimics how a particle, e.g. an electron, would interact with all the different detector layers, and stores the response of the detector in the same format as the actual detector data.

At this point the simulated data and the real data are very similar, but are stored in a raw format, containing detector hits and energy depositions, rather than physics objects. This format is hard to use for further analysis. The final step will thus be to perform the event reconstruction. The purpose of event reconstruction is to convert the raw detector information, be it real or simulated, into physical objects, such as electrons, muons, photons, charged or

neutral hadrons. Each of those objects comes with a set of defining variables, which can be very basic (e.g. p_{T}, η or ϕ) or more complex (e.g. shower shape). The different algorithms and techniques that are used within CMS for this purpose are detailed in Section 6.4.

6.1 What is an "event"?

At the LHC we define an *event* as everything that happens in a proton bunch crossing. These high energy collisions are very complex, often resulting in the production of many hundreds of particles. An illustration of this complexity is shown in Fig. 6.1. A proper description of what happens is impeded by the composite nature of the proton, and by the strong coupling constant of QCD, the quantum field theory governing hadron interactions. Fortunately, it turns out that the full process can be factorized into independent subprocesses, each taking place at different energy scales [77].

The process that is usually of most interest is the interaction between the constituents of the two protons that results in high p_{T} particles. This is referred to as the *hard interaction*. Not every collision produces very hard particles, sometimes protons merely undergo elastic collisions, resulting in very soft scattering products that do not pass the detection thresholds. In general, any interaction producing some detectable particles is called a *minimum bias interaction* [79].

The initial momentum distribution of the partons involved in the hard interaction is contained within *parton distribution functions* (PDFs) describing the structure of the proton. Apart from the hard interaction, the other constituents of the proton can also interact. This usually results in a spray of softer particles, the *underlying event* (UE). Any high momentum particle involved in the collision will emit additional hard QCD radiation. Radiation from particles before the hard interaction is called initial-state-radiation (ISR), whereas radiation off particles produced in the collision is called final-state-radiation (FSR).

Quarks and gluons produced in the collision cannot stay free, they must hadronize in a time scale of $\mathcal{O}(10^{-23}\,\mathrm{s})$. These hadrons, in addition to possible produced leptons, will then pass through the experiment where they can be detected, and used to find out what happened in the collision itself. A complication for the physicists analyzing the data arises from the very high instantaneous luminosity at the LHC. During one bunch crossing there are usually up to 20 pp interactions, collectively referred to as *pileup*. Most of these interactions produce relatively soft particles, but they do add to the overall hadronic activity in an event, and can obscure the interesting hard process. An example of how an event might look like in the CMS detector is shown in Fig. 6.2.

In the next subsections I will elaborate on how to describe an event in a more mathematical way, starting from the factorization theorem. These sections are largely based on Refs. [77, 80–83].

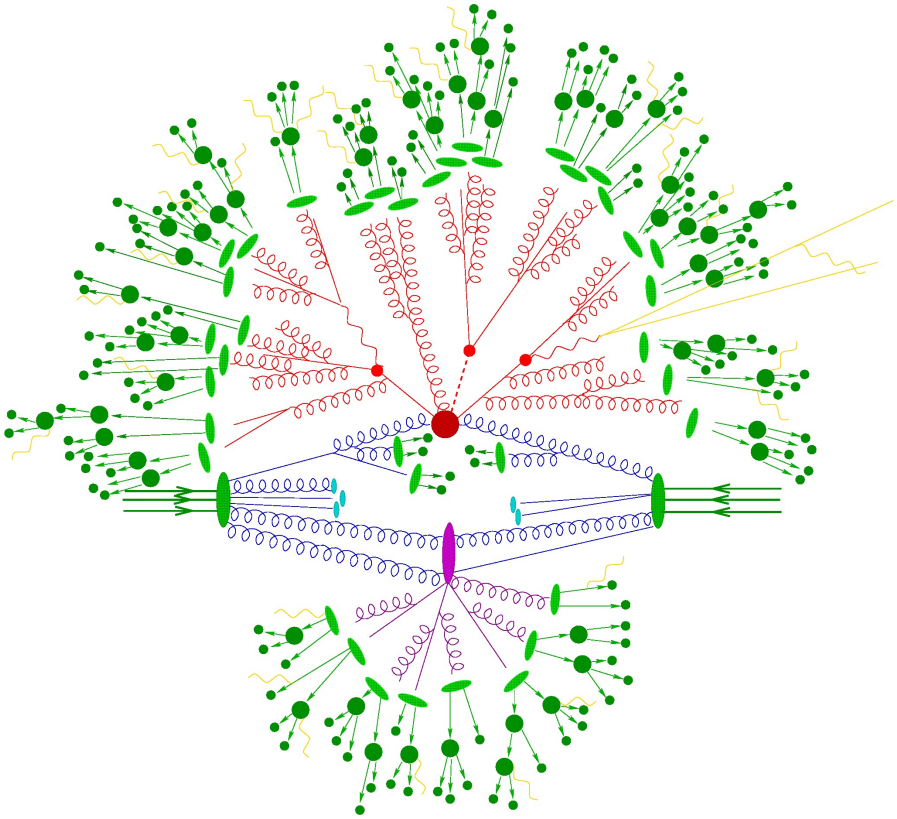

Figure 6.1: Pictorial representation of a *pp* collision event. The hard interaction (big red blob) is followed by the decay of the produced particles (small red blobs). Additional hard QCD radiation is produced (red) and a secondary interaction takes place (purple blob) before the final-state partons hadronize (light green blobs) and hadrons decay (dark green blobs). Photon radiation occurs at any stage (yellow). Figure taken from Ref. [78]

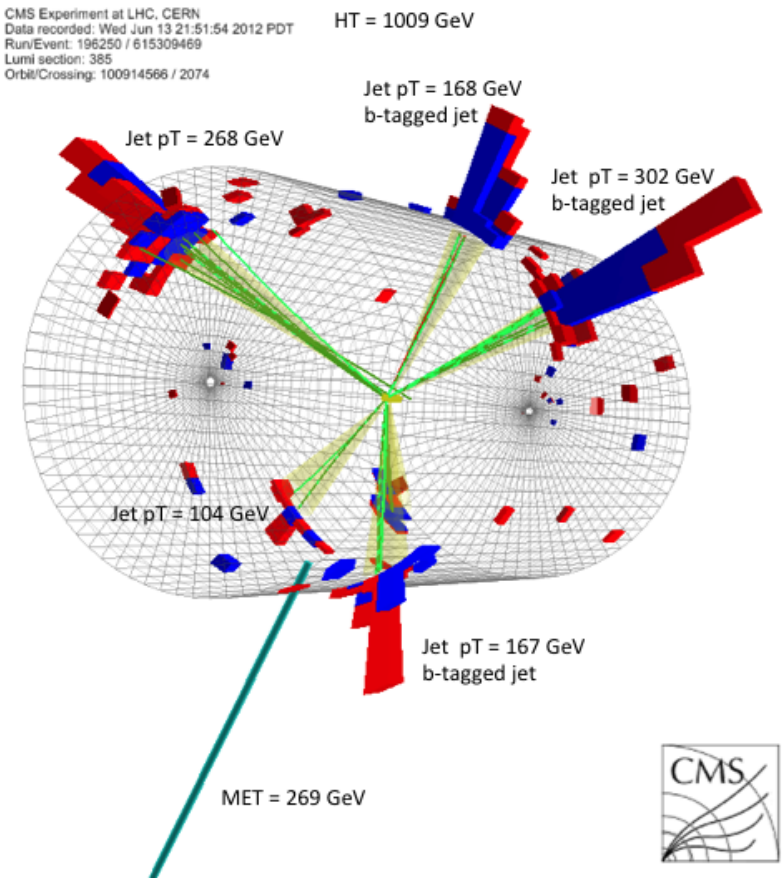

Figure 6.2: CMS event display showing five high p_{T} jets, three of which are tagged as coming from a b quark. Figure from [84].

6.1.1 Factorization theorems

The basic problem addressed by factorization theorems [85] is how to calculate cross sections for high energy processes. In general, these cross sections are a combination of short- and long-distance contributions, and are thus not computable directly in QCD perturbation theory. Factorization theorems allow us to derive predictions for these cross sections, by separating (factorizing) long-distance from short-distance effects. The non-perturbative long-range effects are encapsulated into the parton distribution functions describing the distribution of partons in a hadron. These functions can be measured experimentally, see Section 6.1.2, and most importantly the same functions can be used for different processes. The short-distance hard-scattering cross section can be calculated with perturbation theory because the QCD coupling strength is small at short distances.

The factorization theorem applied to the cross section σ of a hard scattering initiated by two hadrons A and B, illustrated on Fig. 6.3, can be expressed in terms of the parton distribution functions f, and partonic cross section $\hat{\sigma}$:

$$\sigma(s;\alpha_S,\mu_F,\mu_R) = \sum_{a,b}\int_0^1 dx_a \int_0^1 dx_b\, f_{a/A}(x_a,\alpha_S,\mu_F)\cdot f_{b/B}(x_b,\alpha_S,\mu_F)\cdot\hat{\sigma}(\hat{s};\alpha_S,\mu_F,\mu_R), \tag{6.1}$$

with $f_{a/A}(x_a,\alpha_S,\mu_F)$ the probability that a parton a inside a hadron A carries a momentum fraction x_a, s the centre-of-mass energy of the collision, and $\hat{s} = sx_ax_b$ the partonic centre-of-mass energy. The strong coupling constant is denoted by α_S, the factorization scale by μ_F and the renormalization scale by μ_R. The factorization scale defines the (arbitrary) boundary between what is viewed as a short-range versus a long-range interaction. The renormalization scale is also an arbitrary scale, which is needed to regulate the divergencies that appear when computing the partonic cross section in a perturbative expansion. Often, the choice $\mu_F = \mu_R$ is made for convenience. The left-hand side of the equation is in reality independent of the arbitrary choices for μ_F and μ_R. When making computations, a dependence can be introduced because we cannot compute the partonic cross sections up to all orders in α_S.

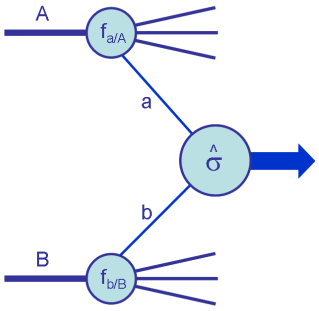

Figure 6.3: Diagram of a hard scattering process, showing the parton distribution functions f and the partonic cross section $\hat{\sigma}$. Figure taken from Ref. [80].

The validity of this factorization theorem can be proven mathematically for certain classes of processes (it is only approximately true for many other processes), but can also be understood intuitively in the context of the parton model. Hadrons are viewed as composite objects, made up of partons held

together by their interactions in a virtual partonic state. Let's consider how a hadron-hadron scattering at high energy and momentum transfer looks like in the centre-of-mass frame. The hadrons appear Lorentz contracted in the direction of the collision, and their internal interactions are time dilated. The higher the centre-of-mass energy, the longer the lifetime of any virtual partonic state will be, and the shorter the time needed for a parton of one hadron to cross the other hadron. At high enough energy, the time needed to traverse the hadron will be much shorter than the lifetime of any partonic state. Each parton inside the hadron can thus be viewed as carrying a definite fraction ξ of the hadron's momentum in the centre-of-mass frame, and so it makes sense to talk about the partons interacting rather than the hadrons. Therefore, the interactions of the partons inside a hadron, which occur at time-dilated time scales before or after the hard scattering, cannot interfere with the interaction of a parton from one hadron with a parton from the other hadron. The cross section for hadron scattering may thus be computed by combining probabilities, rather than amplitudes, and factorization is reached.

6.1.2 Parton distribution functions

A key ingredient to the computation of any cross section at the LHC, is the set of parton distribution functions describing the structure of the proton, as is visible from Eq. 6.1. Physically, PDFs express the fact that hadrons are composite objects, with a time-dependent structure. The PDFs themselves are not physical observables, but rather a more fundamental quantity derived from the actual physical observables such as structure functions, which can be measured in e.g. deep-inelastic scattering processes. Parton distribution functions can be extracted from this data, but only within a specific factorization scheme, order by order in perturbation theory. At leading order they have a very simple physical interpretation: if the PDF for a given particle species p is given by $p(x, Q^2)$, then $p(x, Q^2)dx$ is the probability that a probe of virtuality Q^2 will find a particle of flavour p inside the proton, with a momentum fraction between x and $x + dx$ of the full proton momentum. At higher orders, the PDFs no longer have a clear probabilistic interpretation.

Parton distribution functions satisfy sum rules, governed by the valence content of the hadrons. For a proton we find for the PDFs of the u, d, and s (anti-)quarks:

$$\int_0^1 dx \left(u(x, Q^2) - \bar{u}(x, Q^2)\right) = 2, \tag{6.2}$$

$$\int_0^1 dx \left(d(x, Q^2) - \bar{d}(x, Q^2)\right) = 1, \tag{6.3}$$

$$\int_0^1 dx \left(s(x, Q^2) - \bar{s}(x, Q^2)\right) = 0, \tag{6.4}$$

while we also need to satisfy that the momentum weighted sum of the PDFs of

all particle species is equal to unity, in order to satisfy momentum conservation,

$$\int_0^1 dx\, x \left(g(x, Q^2) + \sum_i [u_i(x, Q^2) + \bar{u}_i(x, Q^2)] \right) = 1, \tag{6.5}$$

where $g(x, Q^2)$ is the gluon PDF, and i runs over all quark flavours.

Looking back to Eq. 6.1, we note that the parton distribution functions depend on the chosen factorization scale. The dependence of the PDFs on the scale Q^2 is described by the DGLAP equations, which can be viewed as renormalization group equations in analogy with the running coupling constant. The DGLAP equations, and thus the PDF evolution, are governed by the so-called splitting functions, P_{ab} , that model the rate for a particle of type a to undergo a collinear splitting to produce a particle of type b.

Parton distribution functions are obtained from global fits to a wide variety of data from many experiments, among which are measurements of deep-inelastic scattering at HERA, and Drell-Yan or inclusive jet production at the Tevatron and the LHC. Since there is only a partial kinematic overlap between this data and the region in (x, Q^2) space where we want to use the PDFs, for example to model the production of supersymmetric particles, the DGLAP evolution is essential for the successful prediction of PDFs in the LHC domain. The splitting functions are now known up to NNLO precision, which reduced the uncertainties on the evolution dramatically, from 30% down to about 2%.

As illustration of the PDF scale dependence, we show in Fig. 6.4 the full set of parton distribution functions for two Q^2 scales, as derived by the NNPDF collaboration [86]. The gluon PDF is seen to dominate for small momentum fractions, and this domination increases as the scale increases. This simply means that as we probe the proton with higher energy, i.e. to smaller length scales, we will find more and more gluons.

Most global analyses, such as the one performed by the CTEQ or MSTW collaborations, use a generic form for the parameterization of the quark and gluon distributions at some reference value Q_0, usually chosen in the range $1 - 2\,\text{GeV}$:

$$f(x, Q_0) = A_0 x^{A_1} (1 - x)^{A_2} P(x; A_3, \ldots). \tag{6.6}$$

The parameter A_1 is associated with small-x behaviour, while A_2 is associated with large x. These two factors are, in general, not sufficient to describe the quark or gluon distribution functions. The term $P(x; A_3, \ldots)$ is a smooth function, depending on one or more parameters, that is introduced to add more flexibility to the PDF parameterization. The various PDF collaborations usually make different choices for the form of $P(x)$. The coefficients A_i are then usually determined by comparing theoretical predictions with the data using the method of least-χ^2 fits. The NNPDF collaboration uses a different approach, and parameterizes $f(x, Q_0)$ by a neural network. Once the PDFs are determined for the reference value Q_0, they are generated for the full (x, Q^2) plane using the DGLAP evolution equations.

Apart from having an estimate for the nominal values of the PDFs in a given kinematic range, it is also important to understand the uncertainties, especially for the gluon PDF, which is the hardest to access experimentally,

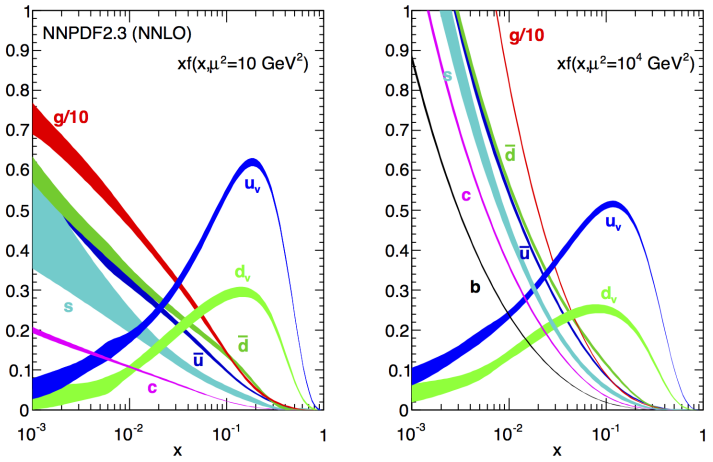

Figure 6.4: Parton distribution functions for different parton species inside the proton for two values for the momentum transfer, as obtained by the NNPDF collaboration [86, 87].

and is constrained mostly by the Q^2 evolution of the quark PDFs. A common method of estimating parton distribution uncertainties is to compare different published parton distributions. This poses a problem since most published PDF sets adopt similar assumptions such that the differences between these sets do not fully capture the uncertainties that actually exist. Several techniques exists that remedy this, and they are used by the PDF collaborations to publish a proper set of uncertainties with each PDF set.

6.1.3 Hard interaction

The partonic scattering cross section describes the hard interaction, and contains all the short-range effects. For the interaction between two partons a and b, resulting in final state F plus anything else (X), it can be written as

$$\hat{\sigma}_{ab\to F+X} = \frac{1}{2\hat{s}_{ab}} |\mathcal{M}_{ab\to F+X}|^2(\Phi_F, \mu_F, \mu_R), \tag{6.7}$$

with $\hat{s} = (p_a + p_b)^2$ the usual Mandelstam variable, and $|\mathcal{M}|^2$ the matrix element squared for the process $ab \to F + X$, appropriately summed and averaged over the relevant helicities and colours. The matrix element depends on the final state phase space Φ_F, and should be evaluated at the factorization scale μ_F and renormalization scale μ_R.

The partonic cross section can be expanded in a perturbative series in the strength of the QCD coupling constant α_S,

$$\hat{\sigma} = \hat{\sigma}_0 + \hat{\sigma}_1 \alpha_S + \hat{\sigma}_2 \alpha_S^2 + \ldots \tag{6.8}$$

The first couple of terms in the perturbative expansion are the terms that so-called *fixed-order predictions* deal with. They are conceptually quite simple;

it is easy to state which contributions are included, and by including further orders in the expansion one can expect to see improvement in the accuracy of the predictions. At leading order we can still compute many inclusive cross sections by hand, although this is often automated, by computing all the relevant tree-level Feynman diagrams and integrating over the appropriate phase space. At next-to-leading order we can distinguish between two sets of extra contributions to the originally considered process: the real emissions resulting in extra quarks or gluons in the final state, and the virtual loops which do not change the number of final state particles, but do impact the cross section.

It is important to note that the complexity of the computations increases mostly with the number of extra loops, rather than the actual order in α_S. Tree-level diagrams can be calculated up to quite high final-state multiplicities, ~ 10, while one-loop diagrams have only been used for processes with up to 3 or sometimes 4 final-state particles, and two-loop diagrams are available only for $2 \to 1$ type processes, such as $pp \to W$. When going to higher orders in the perturbative series, it also becomes more and more tricky to properly combine, i.e. cancel, divergencies between 2-loops, 1-loop and tree-level diagrams. Examples of tree-level diagrams that become divergent is anything produced in association with extra quarks or gluons which could become soft or collinear. These divergencies must be cancelled by the corresponding loop divergencies, otherwise unitarity is violated. In practice this is not always easy to do, especially when experimental cuts need to be applied. The standard technique to deal with this issue is through a *subtraction procedure* which introduces suitable counterterms, adding them to the real diagrams, and subtracting them from the loops, hereby removing the divergencies from the calculation.

Even though the switch from LO to NLO predictions introduces some technical complications, it is still worthwhile to do so, where possible, because of the reduced uncertainties. At NLO, the dependence on the factorization and renormalization scales is much smaller, as this relies on the missing higher order terms, which for NLO contain an extra factor α_S, and are thus smaller.

The strength of the NLO correction is often encapsulated in a so-called NLO k-factor, which is defined as the ratio of the NLO cross section to the LO cross section. K-factors for many processes can be as large as 1.5, much larger than the 10% effect one would expect from considering only the extra factor of α_S. The reason for this is that the terms accompanying that factor of α_S can be quite large. The calculated k-factors can vary for different kinematic regimes within the same process, so care needs to be taken when attempting to scale a LO cross section obtained for some particular corner of phase space.

As explained in the introduction of this chapter, we also need to generate full events for which we can simulate the detector response, rather than only computing inclusive cross sections. This chain often starts by generating events for the hard process only, of course taking into account the parton distribution functions as well. Until recently, event generators based on the perturbative calculation of matrix elements could only generate events up to LO. With the release of MG5_aMC@NLO [88], the automated generation of events at NLO precision is now possible for almost any Standard Model process. More details on how this is done in practice are presented in Section 6.2.1.

6.2 Event generation

In this section I will explain in more detail the various techniques employed by the most common event generators. In particular, I will focus on MadGraph [88, 89] and pythia [90], the programs that generated the events for most of the processes used in the Razor Boost analysis, presented in Chapter 7. The following sections are based on Refs. [77, 80, 82, 88–91].

6.2.1 Matrix element generators

As explained previously, the hard interaction can be described using matrix elements, which can be computed, at least in principle, order by order using perturbation theory. There are a variety of programs available that calculate the tree-level diagrams numerically, and integrate over the relevant phase space. The most widely used are MadGraph, now merged into MG_aMC@NLO, and Alpgen [92]. There are no inherent limits to the number of final state particles that could be produced with these programs, although, in practice, the computation is limited by the factorial growth of the number of diagrams as we go higher in multiplicity of final state particles. In this section I will explain the basic algorithms used by MadGraph to generate events at leading order accuracy. For all details, and a discussion on how event generation at next-to-leading order precision is done, I refer to Refs. [88, 89].

MadGraph allows for automatic generation of matrix elements for collider physics processes, such as decays and $2 \to n$ scatterings. As the user, one first specifies the desired process in terms of initial and final state particles. It is possible to exclude or require the presence of s-channel resonances, and one can force a particular decay chain. Once the process is fully specified, MadGraph computes all Feynman diagrams that can contribute, and writes process-specific code to compute the matrix elements. The user is not restricted to models implemented by default in MadGraph. Feynman rules for any (new) physics model can be obtained via FeynRules [93] and passed to MadGraph via the standardized UFO format [94].

The algorithm to determine all relevant diagrams recursively creates sub-diagrams by merging legs. It can be most easily explained by considering a simple example. We will go through the different steps for the diagram generation of $e^+e^- \to \mathrm{u\bar{u}g}$. The relevant vertices in the Standard Model are $(e^+e^-\gamma)$, $(e^+e^-\mathrm{Z})$, $(\mathrm{u\bar{u}}\gamma)$, $(\mathrm{u\bar{u}Z})$, and $(\mathrm{u\bar{u}g})$. Before the start of the algorithm, the initial state particles are flipped such that only outgoing particles are present. We then proceed as follows.

1. First it is checked whether there is a vertex including all particles. In this example this is not the case.

2. Then all possible two-particle groupings are performed, and the groups are replaced by a single particle according to the allowed vertices. An example is the grouping $(e^+e^-)\mathrm{u\bar{u}g}$, which results in the replacements $(\gamma)\mathrm{u\bar{u}g}$ and $(\mathrm{Z})\mathrm{u\bar{u}g}$. The full list of groupings and replacements for this example is shown in Table 6.1. Each option gets assigned a number according to how many groups were replaced, here either 1 or 2.

Table 6.1: Steps for the diagram generation algorithm employed in MadGraph. Table taken from Ref. [89].

First iteration	Groupings	Replacements
$e^-, e^+, \mathrm{u}, \overline{\mathrm{u}}, \mathrm{g}$	$(e^-, e^+), \mathrm{u}, \overline{\mathrm{u}}, \mathrm{g}$	$(\gamma), \mathrm{u}, \overline{\mathrm{u}}, \mathrm{g}$ $(\mathrm{Z}), \mathrm{u}, \overline{\mathrm{u}}, \mathrm{g}$
	$e^-, e^+, (\mathrm{u}, \overline{\mathrm{u}}), \mathrm{g}$	$e^-, e^+, (\gamma), \mathrm{g}$ $e^-, e^+, (\mathrm{Z}), \mathrm{g}$ $e^-, e^+, (\mathrm{g}), \mathrm{g}$
	$e^-, e^+, (\mathrm{u}, \mathrm{g}), \overline{\mathrm{u}}$	$e^-, e^+, (\mathrm{u}), \overline{\mathrm{u}}$
	$e^-, e^+, \mathrm{u}, (\overline{\mathrm{u}}, \mathrm{g})$	$e^-, e^+, \mathrm{u}, (\overline{\mathrm{u}})$
	$(e^-, e^+), (\mathrm{u}, \overline{\mathrm{u}}), \mathrm{g}$	$(\gamma), (\gamma), \mathrm{g}$ $(\gamma), (\mathrm{Z}), \mathrm{g}$ $(\gamma), (\mathrm{g}), \mathrm{g}$ $(\mathrm{Z}), (\gamma), \mathrm{g}$ $(\mathrm{Z}), (\mathrm{Z}), \mathrm{g}$ $(\mathrm{Z}), (\mathrm{g}), \mathrm{g}$
	$(e^-, e^+), (\mathrm{u}, \mathrm{g}), \overline{\mathrm{u}}$	$(\gamma), (\mathrm{u}), \overline{\mathrm{u}}$ $(\mathrm{Z}), (\mathrm{u}), \overline{\mathrm{u}}$
	$(e^-, e^+), \mathrm{u}, (\mathrm{g}, \overline{\mathrm{u}})$	$(\gamma), \mathrm{u}, (\overline{\mathrm{u}})$ $(\mathrm{Z}), \mathrm{u}, (\overline{\mathrm{u}})$

3. All combinations after the replacement for which fewer than two groupings were replaced, i.e. the first seven in the table, are discarded because they cannot give rise to valid diagrams, or would lead to double counting if grouped further.

4. For the remaining combinations the presence of a valid vertex is checked. Only the final four options have a valid vertex in this example. These diagrams are thus added to the list of possible diagrams.

5. The iteration ends here because any further grouping of these valid diagrams would result in a state that contained less than two replacements. The four diagrams that were generated are shown in Fig. 6.5.

The computation of the squared matrix element for a given process is done via calls to helicity wavefunctions and amplitudes. Helicity amplitudes work on the amplitude level, in contrast to the methods using contraction of Lorentz-indices that work on squared amplitudes. A big advantage is that the complexity of the calculation grows linearly rather than quadratically, and that diagrams are factorized such that the subcomponents, i.e. the helicity wavefunction calls, can be reused between diagrams. A helicity wavefunction is first generated for each external leg in any diagram using the ALOHA [95] package.

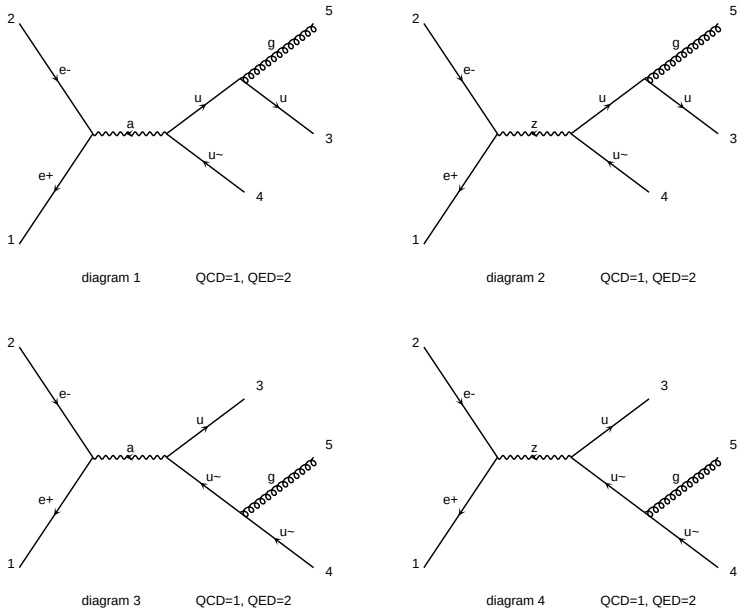

Figure 6.5: Feynman diagrams for the process $e^+e^- \to \mathrm{u\overline{u}g}$

These wavefunctions are then combined into new wavefunctions corresponding to the propagators in the diagram by successive helicity wavefunction calls. The final vertex then corresponds to a helicity amplitude call which returns the value of the amplitude corresponding to this particular diagram.

Particle decays are treated in the same way as the production, allowing for efficient treatment of multiprocesses with the same decay pattern. An example of this is the process $pp \to W^+$, with $W^+ \to \ell^+\nu_l$. This process contains the production processes, $\mathrm{u\overline{d}} \to W^+$, $\mathrm{\overline{d}u} \to W^+$, $\mathrm{c\overline{s}} \to W^+$, $\mathrm{\overline{s}c} \to W^+$, and the decay processes $W^+ \to e^+\nu_e$, $W^+ \to \mu^+\nu_\mu$, $W^+ \to \tau^+\nu_\tau$. All these building blocks only need to be generated once, and can then be combined in all possible ways to obtain the full matrix element.

Once the code for the process under consideration is generated, we can start generating events. MadGraph uses a so-called run_card as configuration for the event generation. In this card the user can specify how many events to produce, which phase space cuts to apply, which parton distribution functions to use, how to choose the renormalization scale, etcetera. Using this configuration, the numerical integration of the matrix element squared over the appropriate phase space is performed, and unweighted events are finally obtained.

The phase space to integrate is usually high-dimensional, and contains many peaks, which are often related to propagators in one of the diagrams becoming large. Efficient sampling techniques are thus critical for the performance of the event generation. Standard MC integration techniques such as importance sampling have the drawback that you need to know a lot about the function f

you wish to integrate in order to find an appropriate, more well-behaved, function g to help the MC integration. Since this is not usually the case for these phase space integrals, the MadGraph program implements a custom integration method, called single-diagram-enhanced multi-channel integration [96]. The method works as follows. Assume that the function to be integrated could be written in terms of a basis of n functions f_i,

$$f = \sum_{i=1}^{n} f_i, \text{ with } f_i > 0, \quad \forall i, \tag{6.9}$$

such that the peak structure of each f_i can be efficiently mapped by a single function g_i. Then, the integration of f reduces to a sum of n independent, and simpler, integrations.

$$I = \int d\Phi f(\Phi) = \sum_{i=1}^{n} \int d\Phi g_i(\Phi) \frac{f_i(\Phi)}{g_i(\Phi)} = \sum_{i=1}^{n} I_i. \tag{6.10}$$

For a generic integration problem, such a basis might be too difficult to identify, but here we can use the physical content of the process and decompose f according to the single Feynman diagrams,

$$f_i = \frac{|A_i|^2}{\sum_i |A_i|^2} |A_{\text{total}}|^2 \tag{6.11}$$

where A_i is the amplitude corresponding to a single Feynman diagram and A_{total} is the total amplitude. Finding the suitable mapping g_i is straightforward, since it can be derived from the known propagator structure of the corresponding Feynman diagram. Since the I_i can be computed independently and then combined, this method is inherently parallel in nature, allowing the use of computer clusters to speed up the computation and thus facilitating the generation of more complicated processes.

Because of its good performance, ease of use, and flexibility, MadGraph is the standard matrix-element generator used by the CMS experiment. Other generators are still used for dedicated processes, or to derive systematic uncertainties on the prediction coming from the details and approximations made by the various event generators.

The result of the event generation is a set of final state, hard particles. Very soft or collinear particles cannot be computed by matrix-element generators, as the matrix element diverges. The next section will cover parton shower programs, whose purpose is exactly to deal with soft and collinear radiation. Section 6.2.3 will discuss a technique on how to match the matrix element computation with the parton shower to obtain the best of both worlds.

6.2.2 Parton shower

The fixed-order matrix-element MC programs discussed in the previous section provide a powerful combination of accuracy and flexibility as long as you want to calculate infrared and collinear safe observables – such as jets, W or Z

bosons, but not pions, kaons, etcetera – and dont need to study regions of phase space that involve disparate physical scales. An example of the latter could be requiring a heavy boson to have a p_{T} much smaller than its mass, leading to large coefficients at all orders in the perturbative expansion. These defects are related to the presence of soft and collinear divergences in the calculations. Real life does not diverge, however. We thus need a different approach to tackle the soft and collinear part of the phase space. This approach is the parton shower.

Parton shower algorithms, such as the one implemented in PYTHIA, describe the evolution in momentum transfer from the high scales associated with the hard process down to the low scales, of order 1 GeV, associated with the confinement of the partons it describes into hadrons. In analogy with bremsstrahlung of photons in QED, a parton (quark or gluon) with high momentum will have some probability to radiate a gluon. This gluon can then radiate more gluons, or it can split in a $q\bar{q}$ pair. This process repeats itself until the energy of the quarks and gluons becomes too low, and hadronization begins. Hadronization is a non-perturbative process, but fortunately it is universal, i.e. it does not depend on the hard interaction, but only on the partons at the low scale after the parton shower.

The probabilities for the various parton splittings are encompassed in the splitting functions, $P_{j\leftarrow i}$, which were already mentioned briefly in Section 6.1.2. Let us first introduce the variable t as

$$t = \ln \frac{Q^2}{\Lambda^2}, \tag{6.12}$$

with Λ the QCD scale. We then find for the differential

$$\mathrm{d}t = \mathrm{d}\ln Q^2 = \frac{\mathrm{d}Q^2}{Q^2}. \tag{6.13}$$

We can view t as a kind of time in the evolution of the parton shower. The smaller t, and thus the lower the scale, the further along in the shower process we are. In terms of the variable t, we can write the differential probability for a parton i to branch into any parton j with momentum fraction z in the following way,

$$\mathrm{d}\mathcal{P}_i = \sum_j \frac{\alpha_S}{2\pi} P_{j\leftarrow i}(z)\mathrm{d}t\mathrm{d}z, \tag{6.14}$$

with the different splitting functions in the collinear limit given by

$$P_{q\leftarrow q}(z) = C_F \frac{1+z^2}{1-z}, \qquad P_{g\leftarrow q}(z) = C_F \frac{1+(1-z)^2}{z}, \tag{6.15}$$

$$P_{g\leftarrow g}(z) = C_A \frac{z^4+1+(1-z)^4}{z(1-z)}, \qquad P_{q\leftarrow g}(z) = T_R(z^2+(1-z)^2), \tag{6.16}$$

where C_F and C_A are colour factors and T_R is a constant depending on the definition of α_S. There are two sets of divergencies that occur in the computation of the branching probability: when the radiated parton becomes extremely soft, or when it becomes collinear with the original parton. The cases where this occurs can in fact not be resolved in any physical measurement. Two exactly

collinear partons look exactly like one parton with the same total momentum. We should thus impose a resolution criterion. Often the chosen criterion is that the relative transverse momentum between the two partons is larger than some cutoff scale Q_0. Imposing this cutoff, then results in a finite resolvable emission probability. Because the total probability of something happening has to be unity, we can find the probability to not have a resolvable emission as one minus the resolvable emission probability. In this way we have avoided computing the divergent pieces, which would have to be added to the divergent loop-correction to the hard process in order to cancel.

Since Eq. 6.14 is a completely general expression that does not depend on the hard process, we can iterate it, using it on a parton resulting from the hard process to generate one branching and then treating the new final state as the hard process, generating another splitting from it, and so on. In what follows we will discuss how this shall be done in practice.

The integral of the branching probability over all allowed z values, according to the particular resolution criterion imposed, and for a given t value, is defined as

$$\mathcal{I}_{j\leftarrow i}(t) = \int dz \frac{\alpha_S}{2\pi} P_{j\leftarrow i}(z) \tag{6.17}$$

The naive probability that a resolved branching occurs during a small range of t values, δt, is given by

$$\sum_j \mathcal{I}_{j\leftarrow i}(t)\delta t, \tag{6.18}$$

where we did not take into account anything which could have happened during the parton shower, before that time. The probability for no resolved emission to occur is then simply given by $1 - \sum_j \mathcal{I}_{j\leftarrow i}(t)\delta t$. If the evolution of parton i starts at $t_{\max}$, then the probability that the parton has not yet branched later in the shower, when $t < t_{\max}$, is given by the product of the probabilities that it did not branch in any of the small intervals δt between t and $t_{\max}$. In other words, letting $\delta t \to 0$, the no-branching probability at time t, given starting point $t_{\max}$, exponentiates, and is given by

$$\mathcal{P}_{\text{no-branching}}(t_{\max}, t) = \exp\left\{ -\int_t^{t_{\max}} dt' \sum_j \mathcal{I}_{j\leftarrow i}(t') \right\}. \tag{6.19}$$

The actual differential probability that the first resolved branching of parton i occurs at 'time' t, which is the actual question we wish to answer, is thus given by

$$\frac{\mathrm{d}\mathcal{P}_i}{\mathrm{d}t} = -\frac{\mathcal{P}_{\text{no-branching}}(t_{\max}, t)}{\mathrm{d}t} \tag{6.20}$$

$$= \left(\sum_j \mathcal{I}_{j\leftarrow i}(t) \right) \exp\left\{ -\int_t^{t_{\max}} dt' \sum_j \mathcal{I}_{j\leftarrow i}(t) \right\}, \tag{6.21}$$

where the first factor in Eq. 6.21 is the naive probability mentioned above, and the second term is an exponential suppression, similar to that found in

the formula for radioactive decay, to account for the fact that if a parton has already branched at t', it can no longer branch at t. This exponential factor, the probability to not branch above a certain scale, here contained in the variable t, is called the Sudakov form factor, $\Delta_i(t_max, t)$.

Implementing this in a Monte Carlo program is conceptually straightforward. First, a random number r is sampled uniformly between 0 and 1. Then the value for t such that $\Delta_i(t_max, t) = r$ is determined. If the solution is above the cutoff t_0, corresponding to the resolution Q_0, then a resolvable branching is generated with scale t, otherwise the shower evolution is terminated. In case a resolvable branching is to be generated, a z value is chosen according to the splitting functions $P_{j\leftarrow i}(z)$, and then the algorithm is started again. Of course, in practice one has to take into account several complications. Different approaches exist for deciding what the initial scale $t_{\max}$ should be. This scale has to match the hard interaction, and could thus be the largest virtuality in the hard scatter, but could also be the centre-of-mass energy. The Sudakov form factor is also not necessarily easily invertible analytically, which can dealt with by using the so-called veto-algorithm. Apart from final state showers, such as explained here, the initial state also undergoes showering. There it is important to properly match the parton shower with the PDF treatment, as well as ensure on-shell partons that take part in the hard interaction. For all details on how this is fully implemented in the PYTHIA shower routine, I refer to the manual [90].

At the end of the parton shower procedure, we end up with many more partons than we had directly after the hard interaction, all which should be described at the low scale, via non-perturbative models. The most widely used hadronization model will be discussed in Section 6.2.4.

6.2.3 Matching the matrix element to the parton shower

On the one hand, parton shower MC programs provide an excellent event description in regions which are dominated by soft and collinear gluon emission, including the hadron-level details that are necessary for the proper simulation of detector effects. On the other hand, matrix element calculations provide a good description of processes where the partons are energetic and widely separated. They also include the effects of interference between amplitudes with the same external partons. The best possible event description can thus only be achieved by combining both approaches. However, the direct addition of the two techniques can lead to double-counting in kinematic regions where the two calculations overlap. This is of particular importance when merging samples for different parton multiplicities, as illustrated in Fig. 6.6. We will thus need a matching between the matrix element and the parton shower to ensure the proper removal of these overlaps. There are several techniques available to perform this matching. I will focus here on the so-called *MLM matching* [97], which is the technique used in CMS to match MADGRAPH with PYTHIA. The MLM technique comprises three main steps, the first of which is done at the matrix element level. Then the partons are showered, and finally the shower

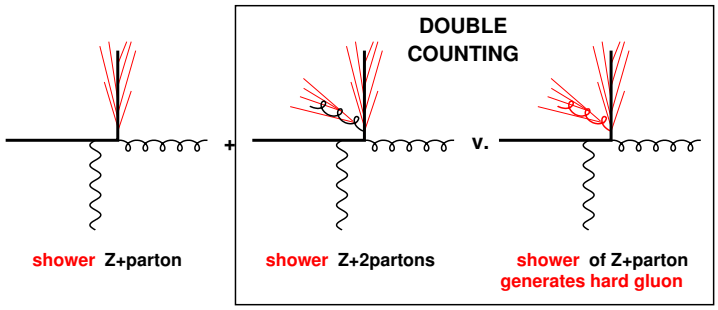

Figure 6.6: Illustration of the double-counting issues that can arise if one naively attempts to shower Z+parton and Z + 2parton events. Partons generated by the matrix element are shown in black, whereas the effects of the parton shower are shown in red. Showering the 1-parton sample could lead to the generation of a hard gluon, which is already included in the matrix-element description of the 2-parton sample, leading to a double-counting. Figure taken from Ref. [82].

jets are matched to the hard partons. In next paragraphs I will discuss each step in more detail.

For each event generated by MadGraph according to the considered hard process, we want to find out how it looks from a parton shower point of view to ensure a smooth transition from the matrix element to the parton shower dominated region. To arrive at this "equivalent parton shower history", we cluster the final state partons. The clustering is performed using the k_{T} jet algorithm, which defines the following two distance measures,

$$k^2_{\mathrm{T},i\mathrm{beam}} = p^2_{\mathrm{T},i} + m^2_i, \tag{6.22}$$

$$k^2_{\mathrm{T},ij} = \Delta R_{ij} \min(p^2_{\mathrm{T},i}, p^2_{\mathrm{T},j}) + \max(m^2_i, m^2_j), \tag{6.23}$$

with $\Delta R^2_{ij} = 2(\cosh \Delta y - \cos \Delta\phi)$. The standard k_{T} clustering starts by finding the smallest of the $k^2_{\mathrm{T},ij}$ or $k^2_{\mathrm{T},i\mathrm{beam}}$, and combining those two partons i and j. The combination then replaces the original two partons, and the clustering is repeated. This continues until there is only a $2 \to 1$ or $2 \to 2$ scattering left. A modification to the standard k_{T} clustering is that only clusterings corresponding to actual Feynman diagrams of the considered model are included. Two quarks of different flavour will thus never be clustered, even if they would have the smallest $k^2_{\mathrm{T},ij}$. Once the clustering is performed, the smallest k_{T} value found must be larger than a chosen cutoff scale, $Q^{\mathrm{ME}}_{\mathrm{cut}}$, otherwise the event is rejected. This cutoff scale is called `xqcut` in the MadGraph configuration files.

In order to mimic the parton shower behaviour, the k_{T} value for each clustering vertex associated with a QCD branching is used as new renormalization scale for α_S in that vertex. This effectively results in an event reweighting. All factorization scales, and the renormalization scale for the hard process, i.e. without additional partons, are constructed by clustering back to the irre-

ducible $2 \to 2$ system, and by using the transverse mass in the resulting frame $\mu^2 = p_{\mathrm{T}}^2 + m^2$. At this point, the events are ready to be transferred to the parton shower. The clustering scales are written in the output file, such that this information is passed along to PYTHIA.

Once the events are passed to PYTHIA, they are showered, using the factorization scale from the previous step as starting point for the shower. Then, before hadronization starts, the showered partons are clustered using the same k_{T} algorithm as before. The resulting jets are required to have a transverse momentum larger than the *matching scale* Q_{match}, with $Q_{\mathrm{match}} > Q_{\mathrm{cut}}^{\mathrm{ME}}$. Partons with a heavy quark as mother are excluded from the clustering.

At this stage, the only missing part is the actual matching between the hard partons, and the jets resulting from clustering the showered partons. Starting from the hardest parton p, we find the closest jet j and declare a match if $k_{\mathrm{T}}(j,p) < Q_{\mathrm{match}}$. We then remove the jet, and repeat for the next hardest parton. If a parton cannot be matched to a jet, we reject the event. If we can match all partons, and there are no additional jets, the event is accepted. In case all partons are matched, but there is an additional jet, then we need to be more careful. If the full process we generated had up to N additional partons at the matrix element level, and the current event had $n < N$ hard partons, then we are in exclusive mode and the event is rejected. The reason is that the additional jet that was added by the parton shower is actually already included at the next multiplicity. In case the current event had N hard partons, we are in inclusive mode, and we will still accept the event as long as the added jet is softer than the softest hard parton in the event. In this way any double-counting is removed. These three cases are illustrated in Figure 6.7.

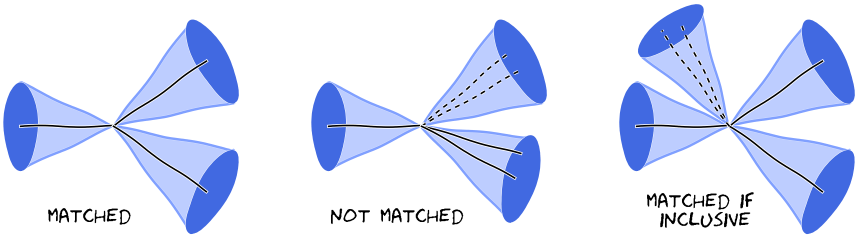

Figure 6.7: Illustration of the matching procedure for an event with three hard partons (solid lines). On the left the jet clustering after the shower results in three jets, that are matched to the three partons. The full event is thus matched, and accepted. In the middle plot the parton shower added two extra partons (dashed lines), and the hard partons were emitted closer together. The jet clustering still results in three jets, but now one hard parton cannot be matched to a jet. The event will thus be rejected. In the right-hand plot each hard parton is matched to a jet, but there is an additional jet present. If we are in inclusive mode, the event is accepted, otherwise it is rejected. Figures adapted from Ref. [98].

It is important to verify that the matching procedure behaves properly. In particular, jet related quantities should have a smooth shape, without any

jumps. The presence of discontinuities indicates that the scales Q_{match} and $Q_{\text{cut}}^{\text{ME}}$ are not chosen correctly to ensure a smooth transition between matrix element and parton shower. The distributions that are typically checked in this regard are the so-called difference jet rates (DJR) [99].

The differential jet rate i is the scale at which a given configuration with $n \geq i$ hard partons passes from being reconstructed as an i-jet one to being reconstructed as an $(i1)$-jet one. In this case, with k_{T} jet clustering, the differential jet rates are simply the actual clustering scales. The $1 \to 0$ differential jet rate (DJR1) is the p_{T} of the last remaining jet after clustering. The $2 \to 1$ differential jet rate (DJR2) is the smallest of the p_{T} of the second last remaining jet and the k_{T} between the second and the first jet, and so on for the other DJR distributions.

Another test of the matching procedure is the stability of the matched cross section when varying the matching scale. In general, the systematic uncertainty associated to the matching procedure can be estimated by varying the matching scale. A common choice is to vary the scale by a factor two up or down.

6.2.4 Hadronization

Real events do not consist of partons but of hadrons. Therefore, the set of post-shower partons must be transformed into a set of primary hadrons, which can then decay further. Hadronization is a non-perturbative transition taking place at the hadronization scale. In the event generation context this scale is by construction identical to the cutoff (resolution) scale of the parton shower. Since we have no idea how to calculate the transition between partons and hadrons from first principles, event generators use QCD-inspired phenomenological models. Although non-perturbative QCD is not solved, we do have some knowledge of the properties that such a solution must have. An important result from lattice QCD calculations is that the potential of the colour dipole field between a charge and an anticharge appears to grow linearly with the separation of the charges, when the separation is greater than about a femtometer. This is known as *linear confinement*, and is used as a starting point for the string model of hadronization. The most widely used model is the Lund model, which is implemented in PYTHIA.

Let us consider the production of a $\text{q}\bar{\text{q}}$ pair. As the quarks move apart, linear confinement implies that a potential

$$V(r) = \kappa r \tag{6.24}$$

is expected at large distances r. This is exactly the potential describing a string with tension κ. We can thus interpret this as a colour flux tube that is being stretched between the quark and the antiquark. From hadron mass spectroscopy the string tension is measured to be about $1\,\text{GeV}/\,\text{fm}$. As the q and $\bar{\text{q}}$ move apart, their kinetic energy is gradually converted to potential energy, stored in the growing string spanned between them. Quark-antiquark fluctuations inside the string field can become real particles by absorbing energy from the string. The original endpoint charges are then screened from each other and the string breaks into two separate colour-singlet pieces, $(\text{q}\bar{\text{q}}) \to (\text{q}\bar{\text{q}}') + (\text{q}'\bar{\text{q}})$.

This process continues until only hadrons remain. Since the string breaks are causally disconnected, they do not have to be considered in any specific time-ordered sequence. In the Lund model, the string breaks are generated starting with the hadrons containing the endpoint quarks, and iterating inwards towards the centre of the string, alternating randomly between the left- and right-hand sides, allowing a single on-shell hadron to be split off in each step. An illustration of the colour flux tube and the breakup of the string system is shown in Fig. 6.8.

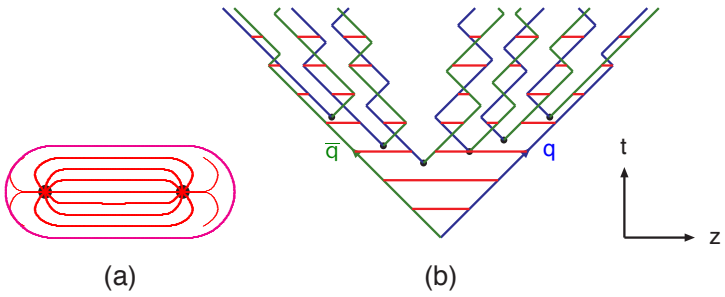

Figure 6.8: (a) A colour flux tube spanned between a quark and an antiquark. (b) The motion and breakup of a string system. Diagonal lines are (anti-)quarks, horizontal lines snapshots of the string field. Figure and caption taken from Ref. [91].

The details of the individual string breaks are not known from first principles. The Lund model uses the idea of quantum mechanical tunnelling, which leads to a flavour-independent Gaussian spectrum for the transverse momentum (w.r.t. the flux tube) of the $q\bar{q}$ pairs. Baryon production can be incorporated by allowing string breaks to occur by the production of pairs of so-called diquarks, loosely bound states of two quarks in a colour antitriplet state. Because the knowledge of hadronization is incomplete, experimental input is needed to tune many of the parameters describing the finer details, such as flavour composition, and the ratio of vector to pseudoscalar mesons.

6.2.5 Generator tuning

Monte Carlo event generators are able to provide a full picture of collider final states, down to the level of individual particles. This allows them to be used as theory reference against which the Standard Model, or a new physics model, can be tested. The accuracy of the prediction depends on the chosen observable, and on the sophistication of the generation. Apart from including higher order corrections, or using better non-perturbative models, it is also crucial to constrain the remaining free parameters of the generation models using existing data. This process is referred to as generator tuning.

Generator models have a vast array of adjustable parameters, but most of these control relatively small details. The few exceptions are the value of α_S in the perturbative domain, and the form of the fragmentation functions

that govern the non-perturbative hadronization process. Tuning all possible parameters is usually done in a highly factorized way, constraining just a few parameters at a time using carefully selected experimental data. Of course, subsequent steps can alter the agreement obtained in the previous steps. Obtaining a full generator tune requires several iterations. Automated tools have been developed in recent years to reduce the amount of manpower required. However, the need for expert input cannot be fully removed.

Different kinds of experimental data are used to constrain different parameters. Tuning of final state radiation and hadronization is mostly done using LEP and other e^+e^- data, which have the advantage that the initial state does not cause extra complications in the interpretation of jet related observables. Constraints on initial state radiation are derived primarily from Drell-Yan events in hadron collisions. While performing the generator tuning it is important to realize that for many observables we do not expect agreement to better than about 5-10%. We should thus take care not to overtune on one single observable, but rather aim for an overall adequate performance.

6.3 Event simulation

Simulation of the CMS detector response is done in one of two ways: a 'full' simulation (FullSim) which is time-consuming but very accurate, or a 'fast' simulation (FastSim), which is much faster, but for which approximations have been made. The following subsections will explain the basic concepts and use cases for both of these options.

6.3.1 CMS Full Simulation using Geant4

The purpose of the CMS full simulation [100–102] is to provide a very accurate description of how particles interact with the CMS detector. These simulated event samples can then be used to understand and demonstrate the power of analysis methods which will later be applied to the real data. They can also be used to derive calibrations, efficiencies and resolutions for high level physics objects in case the available data are not sufficient

The inputs to the detector simulation are the hadronized particles from the event generator. These particles and their four-vectors are passed to the simulation software in the HepMC format [103]. The particles are then propagated through the CMS detector using the Geant4 toolkit [104]. Geant4 contains a large collection of electromagnetic and hadronic physics processes describing the interaction of particles with material, and the resulting energy loss. Examples of this are bremsstrahlung, photon conversion, nuclear interactions, multiple scattering, and showering. In case new particles are produced in these interactions, they are also propagated further.

A key component for the success of the simulation is the precise implementation of the detector geometry, material budget, and magnetic field strength. Not only the active layers need to be accounted for, but also the cooling, cabling, and support structures. Choices must be made for the level of detail to include in the simulation geometry in order to optimize computation speed

versus the correctness of the simulation. Within the CMS software framework there is a single, common implementation of the full detector geometry for both the simulation and the reconstruction, ensuring full compatibility.

The accuracy of the detector implementation has been thoroughly tested with dedicated test beam data, cosmic muon data, and early LHC collision data. Overall, very good agreement was found, and adjustments were made where necessary.

Pileup interactions are also added at this stage of the processing chain. A library of simulated hits of minimum bias events is prepared beforehand, and then used to overlay a number of extra interactions onto the signal event according to a specified pileup scenario. As the bunch crossing time is shorter than the time needed for particles of a given event to fully traverse the CMS detector, we also need to take into account *out-of-time pileup*, which is the effect of previous or subsequent bunch-crossings on the current event. Out-of-time pileup is modelled by modifying the timing of the detector hits when overlaying a minimum bias interaction.

The next step is the conversion of all energy depositions, from signal and pileup events, in the sensitive detector volumes to electronic signals. Electronic noise is also included during this *digitization* step. The output of the digitization is simulated data in a format identical to that of real collision data read directly from the detector. The L1 and HLT decisions, and the objects used to arrive at them, are also included in the simulated data. From this point onwards the simulated events go through the same reconstruction steps as the real data, as will be explained in Section 6.4.

6.3.2 CMS Fast Simulation

The CMS fast simulation [105,106] is a faster alternative to the full simulation explained above. It is intended to be used for physics analyses that require the generation of many samples to span a wide phase-space region, e.g. the vast SUSY simplified model scans. A set of approximations is made, resulting in a speed increase of about a factor twenty.

The interactions simulated in FastSim are electron bremsstrahlung, photon conversion, charged particle energy loss by ionization, charged particle multiple scattering, nuclear interactions, and electron, photon and hadron showering. The various CMS subsystems are modelled in different ways, with various levels of approximation.

The tracking detector is the most complicated CMS subsystem, and the one for which the most approximations are made, including for the geometry. The tracker is modelled by 30 thin nested cylinders, and is assumed to be made of pure silicon, uniformly distributed over each layer. The thickness of each layer in terms of interaction lengths was tuned to reproduce the number of bremsstrahlung photons above a certain threshold as observed in FullSim. Charged particles are propagated between two detector surfaces according to the magnetic field, and experience multiple scattering and energy loss by ionization. The intersections between the trajectories and each tracker layer define the position of the simulated hits that are then converted into reconstructed

hits with a certain efficiency, which is determined from FullSim.

The showers of electrons and photons which hit the electromagnetic calorimeter are simulated as if the ECAL were a homogeneous medium. This is a reasonable approximation because the ECAL crystals are organized to have almost no gaps in between them. Electrons and photons at rapidity values not covered by the electromagnetic calorimeter ($|\eta| > 3$) are propagated directly to the forward hadron calorimeter.

Charged and neutral hadrons are propagated to the start of the ECAL, HCAL and HF after their interactions with the tracker layers. Their energy response is derived from the full simulation. First the energy is smeared according to energy resolution measured in FullSim for single pions. Then, this smeared energy is distributed in the calorimeters using parameterized longitudinal shower profiles.

Muons propagate through the tracker, the calorimeters, the solenoid, and the muon chambers. Both muons coming directly from the main interaction vertex and those produced inside the tracker from the decay of another particle are included. The calorimeter response is treated similarly to that of charged hadrons. In the muon systems the only processes that are taken into account are multiple scattering and energy loss by ionization.

The reconstruction of FastSim tracker hits into tracks is done in a different, faster, way compared to FullSim or data. The truth information is used to group hits together that come from the same particle. Consequently, FastSim does not have fake tracks. Two hits in the same place are also not merged to be one reconstructed hit, as would be the case for the real detector readout. Despite these simplifications, good agreement between FastSim and FullSim track reconstruction is observed.

6.4 Event reconstruction

The CMS event reconstruction aims to provide a global event description in terms of electrons, muons, photons, charged hadrons, and neutral hadrons. Information from all subdetectors is combined to achieve a fully consistent picture of the event. The algorithm that implements this, is called particle flow (PF), and is documented in Refs. [107, 108].

Particle flow relies heavily on the high resolution silicon tracker, and high granularity ECAL. The basic building blocks are tracks, constructed using a very efficient tracking algorithm, and clusters of calorimeter energy deposits. These two elements are then linked together using a linking algorithm. The actual particle flow algorithm uses the links to create a list of reconstructed and identified particles, that are subsequently used for physics analysis. A description of each of these steps will be given in Section 6.4.1.

The particles that are reconstructed by the PF algorithm can be further refined to suit the needs of individual analyses or analysis groups. More stringent identification criteria are usually required, so that the misidentification rate, and thus background rate, is substantially reduced. All physics objects that will be used in the razor boost analysis will be listed in Section 6.4.2.

Reconstructed events are sometimes affected by spurious detector noise or reconstruction failures, leading to anomalous amounts of missing transverse momentum, or to very high p_{T} jets. These events are filtered out by various targeted cleaning algorithms, as discussed in Section 6.4.3.

The PF event reconstruction is applied in the same way to data and simulation, resulting in an overall very good agreement between them. Some quantities cannot be adequately modelled in simulation, however. Event reweighting techniques are employed to account for discrepancies that might influence a physics analysis. In Section 6.4.4 I will discuss the standard event reweighting techniques that are applied in the razor boost analysis.

6.4.1 Particle flow

The particle flow method reconstructs particles, the PF candidates, by combining information from the inner tracker, the calorimeters, and the muon system. Each PF candidate is assigned to one of five object categories: muons, electrons, photons, charged hadrons, and neutral hadrons.

By simultaneously using information from all subdetectors, overlaps between object collections are removed. In a calorimeter only reconstruction, for example, photons and electrons are also reconstructed as jets. Since about 65% of the jet energy is carried by charged hadrons, inclusion of the tracker information in the jet reconstruction results in a much improved energy resolution. For jets originating from a b quark this is even more striking, as energy carried away by muons from the b decay can be included in the jet when taking a PF approach.

The separate components of the PF algorithm are explained below, including a discussion on how particle flow is used to suppress pileup effects.

Iterative tracking

The tracker provides a very good momentum resolution for charged hadrons, better than the calorimeters up to a p_{T} of several hundred GeV. It also gives a precise measurement of the direction of charged particles. For these reasons, the tracker is the cornerstone of the PF algorithm. A tracking efficiency as close to 100% as possible, while keeping the tracking fake rate as low as possible, is thus of the utmost importance. An iterative, Kalman filter based, tracking strategy [109] is used to achieve this.

The track finding algorithm starts by requiring very tight criteria for the track seeds and reconstruction quality, leading to a moderate tracking efficiency, but with a negligibly small fake rate. The hits assigned to those tracks are then removed, and the tracking cycle is repeated two times more for the remaining hits with progressively looser track seeding criteria. The looser criteria increase the tracking efficiency, and the fake rate is kept low because of the reduced combinatorics resulting from the removal of hits in the previous iteration. During three more iterations, the constraints on the origin vertex are also relaxed, allowing the reconstruction of tracks associated with secondary vertices.

With the iterative tracking technique, charged particles with a p_T as low as 150 MeV, as little as three hits, and originating more than 50 cm from the beam axis, can be reconstructed with a fake rate of the order of 1%.

Calorimeter clustering

The calorimeter clustering in the PF method aims for a high detection efficiency and a separation of nearby energy deposits. The calorimeters are also solely responsible for providing a measurement of the energy and direction of neutral hadrons and photons, and should provide additional information on charged hadrons in case the track parameters could not be determined with high precision. The clustering is done separately for the ECAL, HCAL and preshower, and separately for barrel and endcaps. The algorithm consists of three steps.

1. Cluster seeds are identified as calorimeter cells with energy deposits above a certain threshold, and larger than their neighbouring cells.

2. Topological clusters are grown from the cluster seeds by joining adjacent cells that pass a chosen minimum energy threshold related to the level of noise in the electronics.

3. PF clusters are constructed from the topological clusters using an iterative procedure. Each cluster seed gives rise to one PF cluster, even when multiple seeds are part of one large topological cluster. If this happens, the energy of each calorimeter cell is shared among all PF clusters according to the distance between cluster and cell. During the different iterations, the PF cluster position is computed as a weighted average of the positions of the cells. As the cluster position changes, the distance between cluster and cell changes, and thus the energy sharing, which prompts the recalculation of the cluster position. This continues until the cluster positions are stable.

Link algorithm

A particle passing through the CMS detector can leave hits in multiple subdetectors, as was illustrated in Fig. 5.6, and will thus most likely give rise to multiple PF elements. There can be a track, one or more calorimeter clusters, or possibly a track in the muon system. The linking algorithm is designed to link together all elements originating from a single particle, thereby removing any possible double counting from different subsystems. The quality of a given link is quantified by the distance between the linked elements.

A link between a charged particle track and a calorimeter cluster is made by extrapolating the track to the calorimeters, at a depth corresponding to the expected shower maximum. The track is linked to a calorimeter cluster if the extrapolated track position falls within the cluster. The link distance is defined as the distance in the (η, ϕ) plane between the extrapolated track position and the cluster position. Calorimeter clusters originating from bremsstrahlung photons emitted by electrons are captured by extrapolating tangents to a track

at each intersection between the track and a tracker layer toward the ECAL. If the extrapolated tangent position is within a cluster, the cluster is linked to the track as well.

Links between calorimeter clusters are made when the position of a cluster in the higher granularity calorimeter falls within the boundaries of a cluster in the less granular calorimeter. The link distance is defined as before.

A charged-particle track in the tracker and a muon track in the muon system are linked when a global fit between the two tracks returns an acceptable χ^2, the value of which is used as link distance. When several of these 'global muons' can be fit using a given muon track and several tracker tracks, only the global muon that returns the smallest χ^2 is retained.

Particle reconstruction and identification

The blocks of linked elements are converted into a set of identified particles by the particle-flow algorithm. The order of the particle reconstruction follows how clean the signature is. Muons are reconstructed first, neutral hadrons last. The full particle reconstruction and identification algorithm proceeds in this way for each block of linked tracks and clusters.

1. Each global muon is added to the list of PF muons if its momentum is compatible with the momentum determined using tracker information only. The track is then removed from the block.

2. The algorithm then proceeds to reconstruct electrons, using a dedicated method [110]. A Gaussian-Sum Filter is used to refit the candidate electron track, taking into account the possible energy loss by bremsstrahlung, and follow its trajectory to the ECAL. Seeds for the time-consuming fitting procedure are chosen only from the subset of tracks that pass certain identification criteria. An electron is fully identified if its track matches with an ECAL cluster, and if it passes a set of tracking and calorimeter requirements. The electron is then added to the PF electron collection, and the associated electron track and ECAL clusters, including those from bremsstrahlung, are removed from the block before further processing.

3. The remaining tracks are subject to tighter quality criteria, namely the relative uncertainty on the transverse momentum should be smaller than the relative calorimeter energy resolution for charged hadrons. The presence of photons and neutral hadrons will be inferred from a detailed comparison of the track momenta and calorimeter energies.

4. For each HCAL cluster all associated charged hadron candidate tracks are found. If a track traverses more than one HCAL cluster, it is assigned to the closest one. The charged hadron candidate tracks associated with a given HCAL cluster are then matched with the ECAL clusters. The closest ECAL cluster they traverse is assigned to the charged hadron candidate. If the track passes through multiple ECAL clusters, those clusters are first ordered by distance. They are added, one by one, to the

charged hadron candidate for as long as the total calorimetric energy is smaller than the momentum of the charged particle track.

5. If the total reconstructed calorimeter energy is significantly smaller than the total charged particle momentum, there is an inconsistency, and a relaxed muon reconstruction is performed. Tracks that fail more stringent track quality criteria are subsequently removed as well.

6. If on the other hand the total track momentum is smaller than the total calorimeter energy, then the remaining tracks are indeed consistent with stemming from charged hadrons. The tracks are thus added to the list of PF charged hadrons, with the track momentum as charged hadron momentum.

7. For cases where the track momentum is compatible with the calorimeter energy, the charged hadron energy and momentum are refit, using both tracker and calorimeter information. This is of particular interest for high p_{T} hadrons, where the calorimeters provide a better resolution.

8. When there is a substantial excess of calorimeter energy compared to the track momentum, photons and possibly neutral hadrons are reconstructed. First, photons are reconstructed from the ECAL clusters. If this cannot account for the full excess, neutral hadrons are reconstructed from the remainder.

9. Finally, ECAL and HCAL clusters without matching tracks are reconstructed as photons and neutral hadrons, respectively.

At the end of the PF sequence, we now have a list of identified particles which can be used to reconstruct jets. The user can specify which particle types are included in the jet reconstruction. By default, isolated muons and electrons are not included.

Pileup mitigation techniques

The presence of pileup causes extra energy deposits and tracks to be overlaid with those of the hard interaction. This results in a degraded resolution, and less clean signatures. Pileup vertices are usually separated in space from the vertex of interest. The very precise tracker system allows these vertices to be reconstructed, and we can use the particle flow framework to mitigate the effect of in-time pileup, using a technique called *charged hadron subtraction* [111].

Contamination from pileup events is reduced by discarding charged hadron PF candidates that are associated to pileup vertices, prior to jet clustering and any further processing. The leading primary vertex of the event is the one with the largest value of $\sum |p_{\mathrm{T}}^{\mathrm{track}}|^2$. The pileup vertices are all other primary vertices for which the number of degrees of freedom (d.o.f.) in the vertex fit is greater than four. Charged hadrons are assigned to a particular vertex according to the compatibility, expressed as the χ^2/d.o.f., of the track with the proto-vertex reconstructed without the currently considered track. If $\chi^2/\mathrm{d.o.f.} < 20$ for a given track-vertex combination, the track is associated to

that vertex. Charged hadron candidates with a track associated to a pileup vertex are removed. All other tracks, even when not associated to a vertex, are retained.

Since charged hadron subtraction relies on the tracker information, it can only be applied within the tracker acceptance, $|\eta| < 2.5$. It has been shown that this technique is successful at removing a large portion of pileup jets, in addition to removing a significant part of the pileup contribution to jets from the hard interaction. This results in an improved energy and angular resolution.

6.4.2 Physics object identification

The event selection is an integral part of any physics analysis. It determines which events are used, and thus what processes contribute to the data sample. This in turn drives how the backgrounds are estimated, what the sensitivity will be, etcetera. An event selection is most easily described in terms of particles, e.g. two electrons, no muons, at least four jets, as this is the closest to how we think about a given process. The particle flow technique described in the previous section is very compatible with this approach, given that it reconstructs a fully consistent set of identified particles out of the detector hits. However, a more thorough selection of the PF objects is needed in order to ensure that their behaviour is understood, and to ensure that the selected events are not dominated by misidentified particles, or detector artefacts. The physics object groups (POG's) within the CMS Collaboration are in charge of providing general recommendations on how to define each object. These recommendations are based on extensive studies, and are applicable for most analyses, thus reducing the workload for the analysis teams. In the following paragraphs all the standard objects that will be used in the razor boost analysis are discussed.

Primary vertices

We require at least one *good* primary vertex to be reconstructed in each event. This vertex should be associated with at least four charged-particle tracks. It should also lie within 24 cm of the origin of the CMS coordinate system along the beam direction, and within 2 cm in the plane transverse to the beam. These requirements, translated to the CMS nomenclature, are summarized in Table 6.2. In case there are multiple good vertices, we choose the vertex with the highest value of $\sum p_{\mathrm{T}}^2$ of associated tracks to be the leading primary vertex in the event. This vertex is taken as a reference to reconstruct the event, e.g. to perform the track subtraction for pileup removal, for which we use the charged hadron subtraction algorithm, as explained before.

Jets

Most analyses are interested primarily in the quarks and gluon produced in the hard interaction, or in the decay of heavy particles, such as top quarks or W bosons. However, through the process of parton showering and hadronization, the few initial quarks and gluons turn into a multitude of hadrons. Hadrons from a given initial quark or gluon can usually be found close together, they

Table 6.2: Vertex selection criteria.

`isFake()`	$= 0$
`ndof()`	> 4
`z()`	$< 24\,\mathrm{cm}$
`position.Rho()`	$< 2\,\mathrm{cm}$

form a *jet*. The proper description of jets, and the jet definitions that are used to reconstruct them, relies on two properties: infrared, and collinear safety [112]. It is important that a jet definition returns the same set of final jets regardless of whether a parton underwent a collinear or soft splitting. If this is not the case, i.e. the jet definition is infrared or collinear unsafe, then one finds that divergencies in the theoretical computation of jet cross sections do not vanish.

A jet definition comprises two parts: the jet algorithm that defines in which order particles are grouped together, and the recombination scheme that defines how to combine the momenta of the to-be-merged particles. For the latter, the most common choice is to simply add the four-vectors of the particles, which then gives rise to massive jets. For the jet algorithm there are many choices. Here I will focus solely on the anti-k_T algorithm [113], which is the default jet algorithm used by CMS. As for most sequential recombination algorithms, one defines distances d_{ij} between particles i and j (or pseudojets if particles have been combined before), and distances d_{iB} between particle i and the beam. The distance measures are in this case given by

$$d_{ij} = \min\left(\frac{1}{p^2_{\mathrm{T,i}}}, \frac{1}{p^2_{\mathrm{T,j}}}\right)\frac{\Delta R^2_{ij}}{R^2}, \tag{6.25}$$

$$d_{iB} = \frac{1}{p^2_{\mathrm{T,i}}}, \tag{6.26}$$

where $\Delta R^2_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and R is a tuneable parameter determining the size of the jets. The rapidity y of a particle is given by,

$$y = \frac{1}{2}\ln\frac{E + p_z}{E - p_z}. \tag{6.27}$$

The jet clustering proceeds by identifying the smallest of all distances. If it is a d_{ij}, we recombine particles i and j, while if it is d_{iB}, we move i from the list of particles to the list of final jets. All distances are then recalculated and the procedure is repeated until no particles are left. The anti-k_T algorithm results in mostly circular jets, reminiscent of the older cone jet algorithms that are no longer used because they are not infrared and collinear safe.

The input to the jet clustering are the PF candidates that pass the charged hadron subtraction. The clustering itself is done with the anti-k_T algorithm with size parameter $R = 0.5$ (AK5), as implemented in FastJet 3.0.1 [114]. We apply the standard loose identification criteria to the resulting jets, as

defined by the requirements listed in Table 6.3. Jets are required to not be composed of only a single component, as this usually indicates that the reconstruction is of poor quality.

Unfortunately, the calorimeter response to incident particles is not uniform. It is, therefore, not straightforward to translate the measured jet energy to the true particle energy, which is what we want to use to do our analysis. A set of jet energy scale corrections – scalings of the jet four-momentum depending on jet p_{T} and η – are applied to both data and simulation in order to achieve a proper mapping to the particle level. Jet energy corrections within CMS are taken care of in a sequential way, each level of correction taking care of a different effect [115, 116]. The uncertainties associated with the jet energy scale corrections need to be taken into account as systematic uncertainty on any analysis result. How this will be done in the razor boost analysis is explained in Section 7.8.1.

First, the residual effect from pileup is removed using the *L1 corrections*. The effect of pileup on a given jet is quantified by the so-called offset, defined as the difference in p_{T} for a reconstructed jet with added pileup and the same jet without pileup. The effects of charged hadrons from in-time pileup have already been largely reduced by the charged hadron subtraction method. The effect of neutral particles and out-of-time pileup is removed at this stage using a slightly modified version of the *jet area method* [117, 118]. This method uses the effective area of the jets, A, multiplied by the average energy density in the event, ρ, to calculate the energy to be subtracted from the jets. Both real and simulated jets are first corrected with a p_{T}, η, and number of primary vertices dependent offset correction determined in simulation. For data events, an additional data/simulation scale factor is derived from ZeroBias data to correct for remaining η dependent discrepancies. Figure 6.9 shows the size of the energy offset for AK5 jets in the central region, before and after the L1 corrections have been applied. A clear reduction of the overall offset is observed. The dependence on the number of pileup interaction has also been reduced, indicating that the L1 corrections behave properly.

Since the simulation of the detector response is very detailed, see Section 6.3, the jet response in the absence of pileup is in fact very well modelled in simulation. The bulk of the jet energy corrections will thus be derived using simulation, and only the residual differences between data and simulation are derived directly from the data.

The second and third level of corrections, the *L2 Relative* and *L3 Absolute corrections*, are designed to make the jet response flat in η and p_{T}, respectively. They are derived from the simulation together. The L2 correction corrects a jet at arbitrary η relative to a jet in the central area ($|\eta| < 1.3$). Once that is done, the jet energy is translated back to the particle level, such that on average the p_{T} of a reconstructed jet matches that of a jet clustered using generator level particles,

$$< p_{\mathrm{T}}(\mathrm{reco})_{\mathrm{corr}} > = < p_{\mathrm{T}}(\mathrm{gen}) > . \tag{6.28}$$

These are the final corrections applied to jets from simulated events. The size of the L2 and L3 corrections as a function of jet η, for three reference p_{T} values, is shown in Fig. 6.10 on the left-hand side.

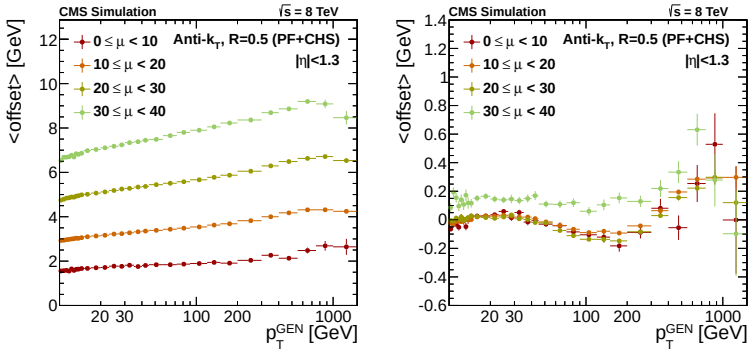

Figure 6.9: The offset shown on the y-axis in these plots is defined as the difference in transverse momentum for a reconstructed jet with added pileup and the same jet without pileup. The left-hand side shows the offset as a function of the generated p_{T} of a jet before the L1 corrections have been applied, and the right-hand side shows the offset after pileup corrections. Different markers represent different levels of pileup. Figures taken from Ref. [119].

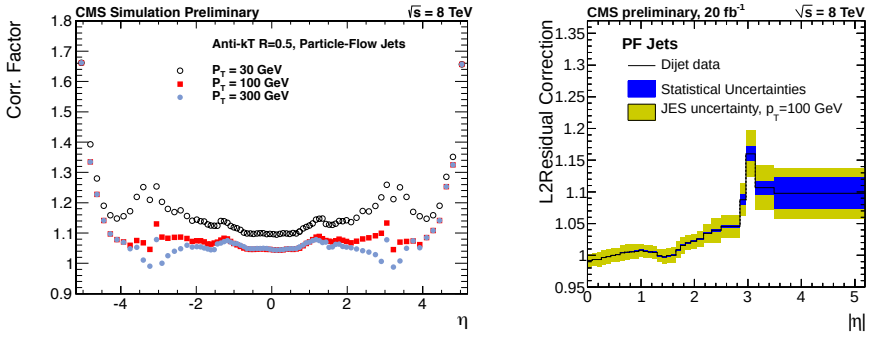

Figure 6.10: [left] The size of the L2 and L3 corrections as a function of jet η for three reference transverse momentum values: 30 GeV (white hollow circles), 100 GeV (red squares) and 300 GeV (blue circles). [right] L2 Residual corrections, obtained from dijet events, as a function of η. The jet energy scale uncertainty is shown with a yellow band, while the statistical uncertainty is shown with a blue band. Figures taken from Ref. [120].

Data events are further corrected by the *L2L3 Residual* jet energy scale *corrections* to take care of the small differences between data and simulation. These corrections are p_T and η dependent, and only correct the relative energy scale. The absolute energy scale was found to be well modelled in the simulation. A dedicated, data-driven approach is employed, using data samples of dijet, γ+jet, and Z+jet events. The L2 Residual correction derived from dijet events is shown on the right in Fig. 6.10.

After all corrections have been applied, jets to be used for analysis are required to have $p_T > 30\,\text{GeV}$ and $|\eta| < 2.4$. The AK5 jets defined in this section will be used for most aspects of the razor boost analysis, except for the reconstruction of boosted hadronic W-candidates. Section 7.4 provides details on the dedicated jet treatment that is used for W tagging.

Table 6.3: Jet selection criteria.

p_T	$> 30\,\text{GeV}$
$\|\eta\|$	< 2.4
`neutralHadronEnergyFraction()`	< 0.99
`neutralEmEnergyFraction()`	< 0.99
`nConstituents()`	> 1
`chargedHadronEnergyFraction()`	> 0
`chargedMultiplicity()`	> 0
`chargedEmEnergyFraction()`	< 0.99

B-tagging

Jets originating from the hadronization of b quarks can be distinguished from other jets, initiated by gluons or light flavour quarks, due to the long lifetime, of the order of $1.5\times10^{-12}\,\text{s}$, of the B hadrons. The non-prompt decay of the B hadrons results in a secondary vertex, displaced by hundreds of micrometers with respect to the primary vertex of the hard interaction. This feature can be used to identify jets as originating from a b quark.

The ability to distinguish b jets is especially important for new physics searches. Many new physics models are associated with production of third generation quarks, whereas this is more rare in the Standard Model. For many searches b jet tagging is an essential tool in suppressing the background from multijet or vector boson production.

CMS has developed several b tagging algorithms [121, 122]. The combined secondary vertex (CSV) algorithm is the most widely used, and combines information on the secondary vertices and the impact parameter of the tracks into a likelihood based discriminant.

The input to any b tag algorithm are high purity tracks, which feature a good track fit with many separate hits, have $p_T > 1\,\text{GeV}$, and lie within the jet cone. The impact parameter of a track is defined as the distance from the

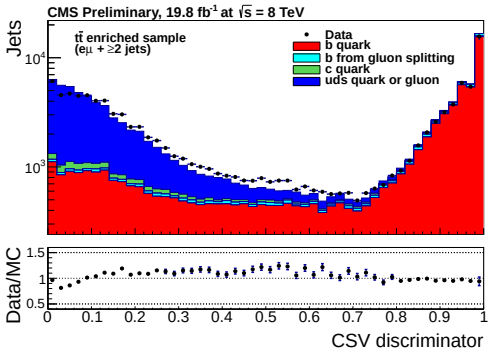

Figure 6.11: Discriminator value for the CSV discriminant for a $t\bar{t}$-enriched data sample.

primary vertex to the track at the point of closest approach. Secondary vertices within jets are reconstructed using an adaptive vertex fitter [123]. The list of secondary vertices is cleaned from candidates that share more than 65% of their tracks with the primary vertex of the event. Vertices that are consistent with originating from a K^0 decay are also removed. Properties of the secondary vertices used in the CSV likelihood discriminant are the flight distance between secondary and primary vertex, the flight distance significance, the vertex invariant mass, and the total p_{T} of all associated tracks. The impact parameter of the tracks is also included. The shape of the CSV discriminant, and a data versus simulation comparison, is shown on Fig. 6.11.

In the razor boost analysis the CSV b tagging algorithm will be applied on the PF jets at two working points [124], which are shown on Table 6.4. The Loose working point (CSVL), corresponding to a misidentification rate of ∼10% and efficiency of ∼85%, will be used to veto b jets, whereas the Medium working point (CSVM), corresponding to a misidentification rate of ∼1% and a typical efficiency of ∼70% , is used to select b jets. The small differences in the discriminant shape, as observed from Fig. 6.11, will result in a slightly different b tagging efficiency in data versus simulation. Scale factors (SF) with associated uncertainties have been derived to correct for this effect, and are provided by the BTAG POG in CMS. Section 7.8.4 will cover how these uncertainties are propagated to the final systematic uncertainties for the razor boost analysis.

Table 6.4: Working points for the combined secondary vertex b jet tagger.

Working point	Discriminator value
Medium	> 0.679
Loose	> 0.244

Muons

The razor boost analysis uses muons that are identified using two different working points, a loose selection and a tight selection, both of which will be detailed below. The loose selection is used throughout most of the analysis, both for vetoing the presence of muons in the signal region, and for selecting single muon events in the control regions. The tight selection is used only to define a control region enriched in $\mathrm{Z} \to \ell\bar{\ell}$ events which is used to derive a systematic uncertainty on the contribution of $\mathrm{Z} \to \nu\bar{\nu}$ in the signal region.

The **loose muon selection** was developed especially for events with a large amount of hadronic activity, where the standard identification criteria were observed to lose efficiency, resulting in less background suppression when vetoing the presence of muons. The details and performance of this optimized selection is documented in Ref. [125]. The main feature is the use of a so-called *directional isolation*. The isolation of a particle is a measure of how far it is from other activity in the detector. The leptons we are interested in, those originating in the hard interaction, are usually separated from other activity, e.g. jets. This is not the case for misidentified muons or for muons from the decay of heavy-flavour jets. Directional isolation is designed to have a better rejection of leptons from these heavy-flavour jet decays, and is defined as

$$\overrightarrow{\mathrm{ISO}}(R) \equiv \sum_{\Delta R_i < R} \delta_i^2 p_{\mathrm{T}\,i}, \tag{6.29}$$

where the sum is over all other particles i within $\Delta R_i < R$ of the muon direction, and δ_i is the angle between particle i and the p_{T}-weighted centroid position (δ_c) of all such particles in (η, ϕ) space. That is, if $\Delta\phi_i$ and $\Delta\eta_i$ are respectively the difference in ϕ and η angles between particle i and the muon, then:

$$\begin{aligned}
\vec{e}_i &\equiv \frac{1}{\sqrt{\Delta\phi_i^2 + \Delta\eta_i^2}} \begin{pmatrix} \Delta\phi_i, \\ \Delta\eta_i \end{pmatrix}, \\
\vec{\delta}_c &= \sum_{\Delta R_i < R} p_{\mathrm{T}\,i} \vec{e}_i, \\
\delta_i &= \angle(\vec{\delta}_c, \vec{e}_i) = \arccos(\vec{\delta}_c \cdot \vec{e}_i / |\vec{\delta}_c|),
\end{aligned}$$

where $\vec{e}_i$ is the unit vector specifying particle i's relative location in (η, ϕ) space with respect to the considered muon, as illustrated in Fig. 6.12. Because of the weighting by δ_i^2, the value for the directional isolation tends to be larger for muons that are near the jet core, e.g. in case of leptonic b decays, compared to the more conventional isolation definition which does not use this weighting.

Apart from the isolation, the identification criteria themselves are also altered from the standard Loose Muon ID from the POG in order to further optimize the muon identification in environments with large hadronic activity. Loose muons are reconstructed using either the global muon algorithm or the tracker-only algorithm. Global muons are required to pass the `GlobalMuonPromptTight` quality criteria, and to have at least two muon chambers containing segments uniquely matched to its inner track. Tracker-only

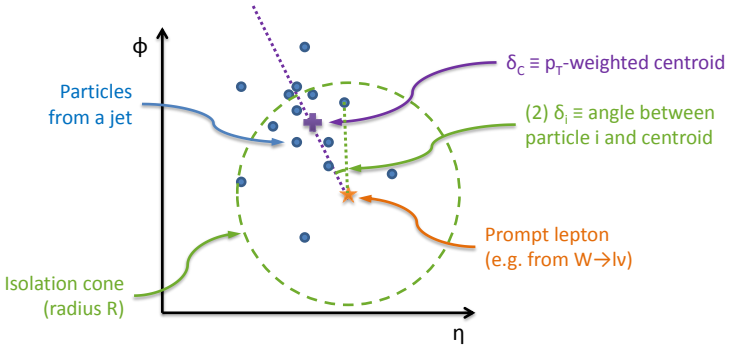

Figure 6.12: Illustration of ingredients used in the computation of directional isolation for a prompt muon, denoted by a star, near some particles from a jet, denoted by points, in the (η, ϕ) plane. For prompt leptons δ_i tends to be small, especially for the high-p_{T} particles near the core of the jet. Figure taken from Ref. [125].

muons are required to pass the `TMLastStationTight` criteria, which require the muon to have compatible hits in the last muon chamber. All selected muons are then required to pass the selection listed in Table 6.5. Some aspects of the selection depend on the muon p_{T} and η; these are summarized in Table 6.6.

The **tight muon selection** follows the recommendation from the Muon POG [126]. In addition to the identification criteria, we also require the tight muon to be isolated. Here we do not use directional isolation, but rather the more standard particle-based relative isolation. This isolation, denoted I_μ, is calculated using the PF candidates in a cone of size $\Delta R = 0.4$ around the muon. Charged-hadron candidates associated with pileup vertices are not taken into account in the calculation of the isolation. However, they are used to estimate the remaining contribution to the isolation coming from neutral hadrons associated with pileup. This contribution is then subtracted. The isolation definition is given by:

$$I_\mu = \frac{I_{Charged} + \max\left(0, I_{Neutral} + I_\gamma - \Delta\beta \cdot I^{PU}_{Charged}\right)}{p^\mu_{\mathrm{T}}}, \tag{6.30}$$

where $I_{Charged}$, $I_{Neutral}$, and I_γ are computed as the sum of the p_{T} of the charged hadrons, neutral hadrons and photons, respectively, in a cone of size $\Delta R = 0.4$ around the muon. The parameter $\Delta\beta$ is set to 0.5, and $I^{PU}_{Charged}$ is the estimated contribution from pileup computed as the sum of the p_{T} of the charged hadrons associated with pileup vertices. The tight muon isolation requirement is $I_\mu < 0.15$. A summary of the tight muon selection can be found in Table 6.7.

Table 6.5: Loose muon definition.

p_{T}	$> 5\,\mathrm{GeV}$
$\lvert\eta\rvert$	< 2.4
`innerTrack().hitPattern().numberOfLostHits()`	≤ 1 if $p_{\mathrm{T}} < 20\,\mathrm{GeV}$ ≤ 4 if $p_{\mathrm{T}} \geq 20\,\mathrm{GeV}$
$\lvert$`innerTrack().dxy(vertex.position())`$\rvert$	p_{T}- and η-dependent
$\lvert$`muonBestTrack().dz(vertex.position())`$\rvert$	p_{T}- and η-dependent
$\overrightarrow{\mathrm{ISO}}(R = 0.2)$	p_{T}- and η-dependent

Table 6.6: Details of the p_{T} dependent thresholds employed in the loose muon selection.

Muon p_{T} (GeV)	d_{xy}(cm) Barrel	d_{xy}(cm) Endcap	d_z(cm) Barrel	d_z(cm) Endcap	$\overrightarrow{\mathrm{ISO}}(0.2)$ Barrel	$\overrightarrow{\mathrm{ISO}}(0.2)$ Endcap
0 - 5	0.052	0.037	0.054	0.076	1.5	2
5 - 10	0.041	0.018	0.042	0.082	3	2.5
10 - 25	0.029	0.013	0.028	0.098	7	7.5
15 - 20	0.014	0.015	0.034	0.1	10.5	9
20 - 40	0.021	0.021	1	0.1	15.5	13.5
40 - 80	0.04	0.2	1	1	32.5	19
80 - 140	0.1	0.2	1	1	54.5	37
140 - 200	0.1	0.2	1	1	87	65.5

Table 6.7: Tight muon definition.

p_{T}	$> 10\,\mathrm{GeV}$
$\lvert\eta\rvert$	< 2.4
`isPFMuon()`	$= 1$
`isGlobalMuon()`	$= 1$
`globalTrack().normalizedChi2()`	< 10
`globalTrack().hitPattern().numberOfValidMuonHits()`	> 0
`track().hitPattern().trackerLayersWithMeasurement()`	> 5
`innerTrack().hitPattern().numberOfValidPixelHits()`	> 0
`numberOfMatchedStations()`	> 1
$\lvert$`innerTrack().dxy(vertex.position())`$\rvert$	$< 0.2\,\mathrm{cm}$
$\lvert$`muonBestTrack().dz(vertex.position())`$\rvert$	$< 0.5\,\mathrm{cm}$
$I_\mu = [$`pfIsolationR04().sumChargedHadronPt()` $+ \max(0.,$ `pfIsolationR04().sumNeutralHadronPt()` $+$ `pfIsolationR04().sumPhotonPt()` $- 0.5 \cdot$ `pfIsolationR04().sumPUPt()`$)$ $] \,/\, p_{\mathrm{T}}$	< 0.15

Electrons

Similar to the muon selection, we identify electrons using two different working points, a loose selection, and a tight selection.

The **loose electron selection** uses directional isolation as described in the previous section, and fully documented in Ref. [125]. A summary of the complete loose electron selection is given in Table 6.8, with the details of the p_{T}- and η-dependent requirements listed in Table 6.9.

The **tight electron selection** is in accordance with the recommendations of the EGamma POG [127]. A summary of the selection can be found in table 6.10. We also require to electron to be isolated. The isolation I_e is calculated using the PF candidates in a cone of size $\Delta R = 0.3$ around the electron, and then corrected with an estimate of the median energy from pileup as calculated with the `FastJet` algorithm in a similar way to the L1 jet corrections explained in Sec. 6.4.2. We require that this corrected isolation, relative to the p_{T} of the electron is less than 0.15.

$$I_e = \frac{I_{Charged} + \max(0, I_{NeutralHad} + I_\gamma - A\rho)}{p_{\mathrm{T}}^e}, \tag{6.31}$$

with A the effective area of the cone, and ρ the average pileup density. Small discrepancies exist between the electron identification efficiency in data and in simulation. Scale factors are provided to correct for this effect, see Section 7.8.6.

Isolated tracks

In order to suppress the decays of both taus and other leptons that do not pass the loose selection, we can veto events for which an isolated track is present [128]. Isolated tracks are selected from the charged PF candidates with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and longitudinal track-primary vertex distance of $d_z < 0.05\,\mathrm{cm}$. They are required to have a relative isolation in a cone of $\Delta R = 0.3$ of less than 0.1. In the razor boost analysis the isolated track veto will only be applied in the hadronic event selections, and not in the control regions which require the presence of a lepton.

Missing transverse momentum

The missing transverse momentum, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, associated with a given event is computed as the negative vector sum of the transverse momentum of all PF candidates i,

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_i \vec{p}_{\mathrm{T}}^{\,i}. \tag{6.32}$$

and its magnitude is denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$.

The corrections to the jet energy scale discussed above are propagated to the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ as well. Within CMS this type of missing transverse momentum is known as type-1 corrected $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ [129].

$$\vec{E}_{\mathrm{T,type\text{-}1}}^{\mathrm{miss}} = \vec{E}_{\mathrm{T,raw}}^{\mathrm{miss}} + \sum_i \vec{p}_{\mathrm{T,raw}}^{\,i} - \sum_i \vec{p}_{\mathrm{T,corr}}^{\,i} - \sum_i \vec{\mathcal{O}}^{\,i} \tag{6.33}$$

Table 6.8: Loose electron definition.

	Condition	Barrel	Endcap
p_{T}		$> 5\,\mathrm{GeV}$	$> 5\,\mathrm{GeV}$
$\lvert\eta\rvert$		< 1.442	$1.556 - 2.5$
`gsfTrack().numberOfLostHits()`	$p_{\mathrm{T}} < 20\,\mathrm{GeV}$	$= 0$	$= 0$
`gsfTrack().hitPattern().numberOfValidPixelHits()`	$p_{\mathrm{T}} < 10\,\mathrm{GeV}$	≥ 2	≥ 1
$\lvert$`gsfTrack().dz(vertex.position())`$\rvert$		p_{T}- and η-dependent	
$\overrightarrow{\mathrm{ISO}_{\mathrm{c}}}(R = 0.3)$, calculated from charged particles only		p_{T}- and η-dependent	
$\overrightarrow{\mathrm{ISO}}(R = 0.2)$, barrel only, calculated using all particles		p_{T}- and η-dependent	

Table 6.9: Details of the p_{T} dependent thresholds employed in the loose electron selection.

Electron p_{T} (GeV)	d_z(cm) Barrel	d_z(cm) Endcap	$\overrightarrow{\mathrm{ISO}_{\mathrm{c}}}(0.3)$ Barrel	$\overrightarrow{\mathrm{ISO}_{\mathrm{c}}}(0.3)$ Endcap	$\overrightarrow{\mathrm{ISO}}(0.2)$ Barrel
0 - 5	0.03	0.09	0.5	0.5	2
5 - 10	0.05	0.09	1.5	2.5	4.25
10 - 25	0.05	0.09	4.5	6.5	8.75
15 - 20	0.05	0.11	7.5	9	11
20 - 40	0.2	1	10	10.5	20.8
40 - 80	1	1	18.5	18.5	200
80 - 140	1	1	44	66.5	200
140 - 200	1	1	81.5	70	200

Table 6.10: Tight electron definition.

	Barrel	Endcap
p_{T}	$> 10\,\mathrm{GeV}$	$> 10\,\mathrm{GeV}$
$\lvert\eta\rvert$	< 1.442	$1.556 - 2.5$
$\lvert$`deltaEtaSuperClusterTrackAtVtx()`$\rvert$	< 0.004	< 0.005
$\lvert$`deltaPhiSuperClusterTrackAtVtx()`$\rvert$	< 0.030	< 0.020
`sigmaIetaIeta()`	< 0.010	< 0.030
`hadronicOverEm()`	< 0.120	< 0.100
1.0/`ecalEnergy() - eSuperClusterOverP()/ecalEnergy()`	< 0.050	< 0.050
`gsfTrack().trackerExpectedHitsInner().numberOfHits()`	≤ 0	≤ 0
`passConversionVeto()`	$= 1$	$= 1$
$\lvert$`innerTrack().dxy(vertex.position())`$\rvert$	$< 0.02\,\mathrm{cm}$	$< 0.02\,\mathrm{cm}$
$\lvert$`gsfTrack().dz(vertex.position())`$\rvert$	$< 0.1\,\mathrm{cm}$	$< 0.1\,\mathrm{cm}$
I_e	< 0.15	< 0.15

Table 6.11: Isolated track selection.

p_{T}	$> 10\,\mathrm{GeV}$
`charge()`	> 0
$d_z(\mathrm{PV}, \mathrm{track})$	$< 0.05\,\mathrm{cm}$
$I_{\mathrm{track}_i} = \frac{\sum_{j \neq i} p_{\mathrm{T}\,j}}{p_{\mathrm{T}\,i}}$	< 0.1

where $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}{}_{\mathrm{raw}}$ is the uncorrected missing transverse energy, $\vec{p}_{\mathrm{T}\,\mathrm{raw}}$ is the uncorrected jet p_{T}, $\vec{p}_{\mathrm{T}\,\mathrm{corr}}$ is the fully corrected jet p_{T}, and $\vec{\mathcal{O}}$ is the average offset due to pileup. Only jets with $\vec{p}^{\,i}_{\mathrm{T,corr}} > 10\,\mathrm{GeV}$ are included in the sum. The average pileup offset underneath jets is included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector sum to ensure that the pileup offset remains isotropic and does not cause any bias.

The missing transverse momentum is sensitive to detector malfunctions and to various reconstruction effects that result in the mismeasurement of particles or their misidentification. Precise calibration of all reconstructed physics objects is thus crucial for the performance of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$.

No explicit selection will be placed on $E_{\mathrm{T}}^{\mathrm{miss}}$ in the razor boost analysis selection, but it is used in the definition of the razor variable R^2, to be introduced in Section 7.3.

6.4.3 Event cleaning

The full CMS data taking and event reconstruction process is very intricate. Every now and then a subdetector might not have behaved properly, or a reconstruction algorithm could have failed. Events affected by such failures need to be removed from the selection, as they can create artificially high missing transverse momentum and would then end up in the signal region of many supersymmetry searches. The following cleaning filters are applied:

- The `EcalDeadCellTriggerPrimitiveFilter`, which removes events where dead cells in the ECAL produce anomalous activity.
- The `hcalLaserEventFilter`, which removes events where the HCAL laser produces anomalous activity.
- The `trackingFailureFilter`, which removes events where the tracking algorithm does not perform properly.
- The `CSCTightHaloFilter`, which removes events contaminated by beam halo.
- The `HBHENoiseFilter`, which removes events featuring large hadronic calorimeter noise.
- The `eeBadScFilter`, which removes events featuring high amplitude anomalous pulses due to bad ECAL super-crystals.

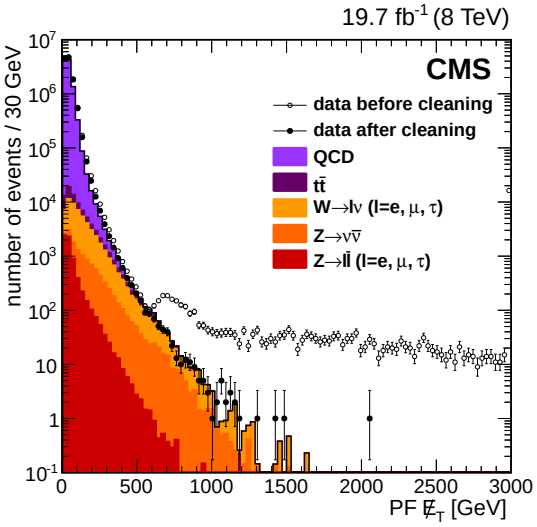

Figure 6.13: The PF $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ distribution for events passing a dijet selection without cleaning filters applied (open markers), with cleaning filters applied (filled markers), and simulated events (filled histograms). Figure taken from Ref. [129].

- The `trkPOGFilters`, which remove events due to track reconstruction anomalies, such as events with partly aborted track reconstruction and events affected by the Strip Tracker coherent noise.
- The `primaryVertexFilter`, which removes events that do not have a good primary vertex.
- The `noscrapingFilter`, which removes events with a large multiplicity of low quality tracks.

More details on these filters can be found in Ref. [130]. The effect on the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is shown in Fig. 6.13. There is a very clear reduction in the $E_{\mathrm{T}}^{\mathrm{miss}}$ tail in data, which brings data and simulation in agreement. The `CSCTightHaloFilter` and `HBHENoiseFilter` filters are not applied for simulated samples that are passed through the fast CMS detector simulation because the necessary input collections are not produced.

HCAL noise filter In addition to the standard filters listed above, we also use an extra cleaning selection designed to remove events with spurious HCAL noise originating in the outer barrel of the HCAL. Energy deposits in the HO are included in the computation of the missing transverse momentum using the particle flow algorithm, but are not included in the missing transverse energy obtained from calorimeter information only. A selection requiring no substantial discrepancy between the two $E_{\mathrm{T}}^{\mathrm{miss}}$ definitions is thus effective at reducing the contribution of these noisy events.

We reject events in which the particle flow missing transverse energy vector, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{PF})$, is flipped with respect to the calorimeter based one, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{Calo})$. To accomplish this we compute the absolute value of the difference in polar angle, $|\Delta\phi_{\mathrm{PF,Calo}}|$, taken in the range $[0, 2\pi[$, and defined as

$$|\Delta\phi_{\mathrm{PF,Calo}}| = \min\left(\phi^{\mathrm{PF}} - \phi^{\mathrm{Calo}}, 2\pi - \phi^{\mathrm{PF}} + \phi^{\mathrm{Calo}}\right), \tag{6.34}$$

with

$$\phi^{\mathrm{PF/Calo}} = \arctan\left(\frac{\vec{E}_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{PF/Calo})|_y}{\vec{E}_{\mathrm{T}}^{\mathrm{miss}}(\mathrm{PF/Calo})|_x}\right). \tag{6.35}$$

Events for which $|\Delta\phi_{\mathrm{PF,Calo}}|$ falls in a 1 radian window centred around π are removed.

$$\left||\Delta\phi_{\mathrm{PF,Calo}}| - \pi\right| < 1 \tag{6.36}$$

6.4.4 Event reweighting

The generation and simulation of events are tuned to mimic the data. However, the complete data taking conditions, in particular the pileup profile, are not fully known before data taking starts. It is thus impossible to mimic the data in all aspects. Furthermore, the event generation itself is also not perfect. Many details of the hadronization process are still unknown, and state of the art event generators can only compute hard physics processes up to maximally NLO precision, whereas data contains all orders. All these effects can lead to discrepancies between the observed data and the simulation.

To correct for some of these imperfections, event reweighting prescriptions have been developed. I have already mentioned some correction factors that should be applied to various objects, such as the jet energy scale corrections for jets. In the next subsections I will cover the reweightings that have to be applied to the full event, rather than a particular object. These include the corrections for mismodelling of the pileup distribution, the initial state radiation, and the top quark p_{T} spectrum for the $t\bar{t}$ simulation.

Pileup reweighting

The distribution of the number of pileup interactions is different in data with respect to simulation. Given that the number of pileup interactions can have an influence on various aspects of the reconstruction, such as the identification of primary vertices or lepton isolation, the simulated events should be reweighted such that their pileup distribution matches that of data [131].

The pileup distribution in data is provided centrally by the Physics Validation Group for each data taking period. This distribution depends on the total pp inelastic cross section. In simulation, the pileup distribution is taken from truth information, through the variable `trueNumInteractions` which stores how many pileup events were overlaid on the hard scatter. The pileup weights are computed as the ratio of the normalized pileup distributions in data and simulation, and should be applied to all simulated events. The distribution of the pileup in data and simulation, and the corresponding pileup weight is

shown in Figure 6.14. As can be seen, the initial guess for the pileup distribution, which was implemented in the simulation, was not perfect, resulting in an effective reduction of the statistical precision of the simulated samples.

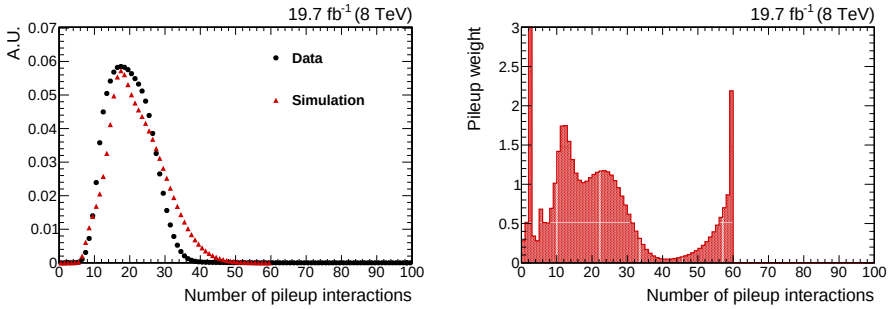

Figure 6.14: [left] Comparison of the distribution of the number of pileup interactions in data and in simulation. [right] Pileup weight as a function of the number of interactions.

ISR reweighting

Searches for new physics often rely on an initial state boost of the produced system in order to have experimental acceptance for the signature under consideration. This is especially important for models featuring a compressed mass spectrum. A high-p_{T} ISR jet can be used to suppress background, or the boost can raise the momentum of jets or leptons in the decay chain to a level that is detectable. A mismodelling of the initial state radiation, or uncertainty on the modelling, will thus directly impact the interpretation of these searches.

A study was performed to investigate how well the ISR is modelled in the simulation by evaluating the agreement between data and simulation in the boost p_{T} for Z+jets and $t\bar{t}$ events [132, 133]. For Z+jets events the boost p_{T} was measured from the leptonic decay products of the Z boson. For $t\bar{t}$ events the ISR radiation was measured using the hadronic recoil system, which is computed from all jets except for the b-tagged jets from the $t\bar{t}$ decay.

It was found that the initial state radiation is not well modelled at high p_{T}, as illustrated in Fig. 6.15. The mismodelling can be corrected by applying a scale factor, with associated uncertainty, which was derived from the observed disagreement. The scale factor depends on the p_{T} of the system recoiling against the ISR jets. This system could be e.g. the $t\bar{t}$ system, the Z boson, or the $\tilde{g}\tilde{g}$ system for a SUSY event. The uncertainty on this scale factor is taken to be the difference between the scale factor and unity. The CMS SUSY group recommends to apply this ISR reweighting to all SUSY signal samples. The prescription is summarized in Table 6.12.

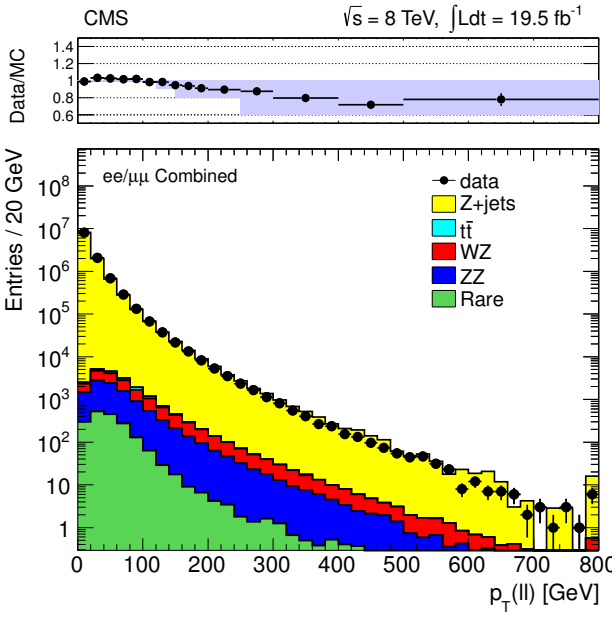

Figure 6.15: Comparison of data to simulation for the p_{T} of the dilepton (ee or $\mu\mu$) system in Z+jets events. The prediction from simulation is normalized to the total data yield to compare the shapes of the distributions. The ratio of data/simulation is shown at the top of the figure, and the light blue band shows the weights derived for simulation and the variation to assess systematic uncertainties. Figure from Ref. [132].

Table 6.12: ISR reweighting prescription.

p_{T} of recoiling system (GeV)	Scale factor
≤ 120	1.00 ± 0.00
$120 - 150$	0.95 ± 0.05
$150 - 250$	0.90 ± 0.10
> 250	0.80 ± 0.20

Top quark p_{T} reweighting

Differential top-quark-pair cross section analyses have shown that the shape of the p_{T} spectrum of top quarks in data is softer than predicted by simulation [134, 135]. To remedy this, events are reweighted based on the p_{T} of the generator level t and $\bar{t}$ quarks in the $t\bar{t}$ simulation. The event weight, w_{TopPt}, is computed as a function of the generated p_{T} of both the top and anti-top quark in the event:

$$w_{\mathrm{TopPt}} = \sqrt{SF_t \cdot SF_{\bar{t}}} \tag{6.37}$$

$$SF(p_{\mathrm{T}}^{gen}) = \exp(a + b\, p_{\mathrm{T}}^{gen}) \tag{6.38}$$

with $a = 0.156$ and $b = -0.00137$. The uncertainty associated with this reweighting is taken to be equal to the full size of the reweighting, which gives for the one standard deviation up and down variations of the event weight:

$$+1\ \sigma : w_{\mathrm{up}} = w_{\mathrm{TopPt}} \cdot w_{\mathrm{TopPt}}, \tag{6.39}$$

$$-1\ \sigma : w_{\mathrm{down}} = 1. \tag{6.40}$$

The effect of this reweighting on the data/simulation agreement in a $t\bar{t}$ enriched region is shown in Fig. 6.16. The x-axis on these plots shows the razor variable $\mathrm{M_R}$, which will be defined in Section 7.3. It is clear that including this reweighting improves the agreement.

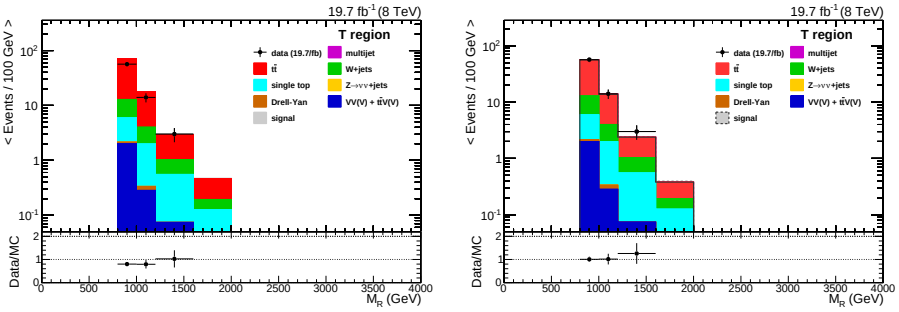

Figure 6.16: Distribution of the razor variable $\mathrm{M_R}$ in data and simulation before applying the top p_{T} reweighting (left), and after (right).

Chapter 7

The razor boost analysis

The razor boost analysis presented in this thesis is a search for new massive particles in hadronic final states with at least one highly boosted W boson and b jet. Boosted W boson candidates are identified using jet substructure techniques. The razor variables, M_R and R^2, are used to characterize the signal as a peak on a falling background.

The search is performed using pp collision data collected by CMS at a centre-of-mass energy of 8 TeV, corresponding to an integrated luminosity of $19.7\,\text{fb}^{-1}$. Backgrounds from Standard Model processes are predicted using a data-driven method. A data control region, binned in M_R and R^2, is defined for each of the main background components. These control regions are used together with global transfer factors, determined in simulation, to obtain the background estimate in the corresponding signal bins on the (M_R,R^2) plane. The validation of the background estimation technique is done using closure tests.

The final results are interpreted in the context of natural SUSY, in terms of top squark production from gluino decays, where the top squarks decay either to a charm or top quark, and a neutralino.

The various steps in the analysis are discussed at length in this chapter. I will first cover the motivation and general strategy of the analysis in Sections 7.1 and 7.2. Then, the *razor variables*, which are our most important discriminating variables, will be derived in Section 7.3. Section 7.4 details the technique used to tag highly boosted W bosons. The datasets and triggers are listed in Section 7.5, followed by the event selection in Section 7.6. The full statistical treatment, with its likelihood based approach, is explained in Section 7.7. In Section 7.8 the different sources of systematic uncertainties are discussed, followed by the results of the full background estimation in Section 7.9. This chapter concludes, in Section 7.10, with the interpretation of the results in terms of several simplified model spectra.

7.1 Motivation

The CERN LHC has provided sufficient data to conduct a large variety of searches for physics beyond the Standard Model. As explained in Chapters 3 and 4, supersymmetry is among the best-motivated candidates for new physics and predicts the existence of supersymmetric partners for each of the SM particles. Scenarios with non-degenerate supersymmetric particle spectra, with cross sections as low as $\sim$1 fb, have been explored in many final states [136]; however, as yet no traces of new physics have been found.

Recently, the focus of searches has turned towards natural SUSY, in which the Higgs boson mass can be stabilized without excessive fine tuning. Natural SUSY requires the existence of a light top squark, $\tilde{t}_1$, and a somewhat light gluino, $\tilde{g}$, while accommodating mass scales for other supersymmetric particles that are beyond the direct reach of current LHC data. More detailed information on natural supersymmetry was already presented in Section 4.6. The possibility that the top squark could be light has motivated several searches by the CMS and ATLAS collaborations [132, 137–143] for the direct production of top squarks. The sensitivity of many of these searches, however, diminishes when the mass of the top squark approaches that of the lightest SUSY particle (LSP), assumed in the remainder of this thesis to be the lightest neutralino, $\tilde{\chi}_1^0$. Searches looking specifically for $\tilde{t}_1 \to t\tilde{\chi}_1^0$ also become less sensitive when the mass difference, Δm, between the top squark and the LSP is comparable to the top quark mass, m_t. These gaps in the sensitivity are illustrated in Fig. 7.1, which shows the general form of the exclusion limits for direct stop production, on the $(m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0})$ plane. Let us now examine the three regions, shown on the figure with coloured ellipses, where general searches lack sensitivity, in order to determine why these regions are hard to probe, and whether we can find a strategy to deal with the issues.

Compressed scenario In general, models with mass spectra featuring small mass splittings are called *compressed scenarios* or *compressed spectra*. This case thus corresponds to the left-most gap in Fig. 7.1, where Δm is very small, smaller than the W boson mass in particular. In this scenario, the top squark decays to the LSP and other soft decay products, either resulting from the loop-induced decay $\tilde{t}_1 \to c\tilde{\chi}_1^0$, or from the four-body decay $\tilde{t}_1 \to \mathrm{b}f\bar{f}\tilde{\chi}_1^0$. These soft decay products, jets and/or leptons, are difficult to detect. They are hard to reconstruct, and when reconstructed they often fall below the p_T thresholds that define the objects. Therefore, in order to be sensitive to such processes, one should not rely on the presence of these objects, but rather on something else, such as the presence of jets from initial state radiation (ISR). Both ATLAS and CMS have performed searches using this technique [144, 145].

Stealthy stop scenario The scenarios where $\Delta m \approx m_t$, are often referred to as *stealthy* scenarios. The reason for this is that when Δm approaches the top mass, the signature of top squark production is very similar to that of Standard Model $t\bar{t}$ production, which has a much higher cross section. Consequently, the signal from direct stop production is hidden underneath a much larger $t\bar{t}$

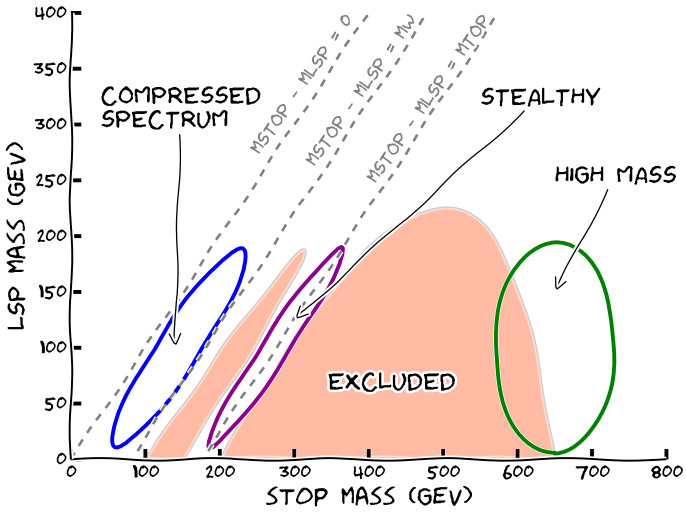

Figure 7.1: General form of the exclusion limits for direct stop production on the $(m_{\tilde{t}_1}, m_{\tilde{\chi}_1^0})$ plane. The red shaded area is the approximate region that has been excluded by a range of early Run 1 searches. The three coloured ellipses indicate the regions that are hard to probe: the compressed spectra, stealthy top squark scenario and high mass top squarks.

background. An alternate way to approach stealthy stops is, for instance, to assume that the heavy top squark $\tilde{t}_2$ is also accessible at the LHC, and decays to the $\tilde{t}_1$ via either a Higgs or Z boson [143]. This results in a longer decay chain, which provides extra handles, such as additional b quarks or leptons.

High mass scenario The last gap that is present in the sensitivity of searches for the direct production of top squark pairs, is the high mass region. In this region the signature is actually very striking, with usually large hadronic activity and/or missing transverse momentum. The problem lies in the rather low expected cross section for direct stop production at 8 TeV, as is clear from Fig. 4.1. We would need much more data than the $20\,\text{fb}^{-1}$ that is available to detect this process at those large masses. Of course, with the restart of the LHC at 13 TeV centre-of-mass energy fast approaching, we can expect to close part of this gap very soon.

Apart from the approaches mentioned above, there is another option to tackle the compressed and stealthy scenarios, namely, looking for top squarks in gluino decays. This is exactly the focus of the razor boost analysis. Specifically, we consider gluino pair production in which the gluino decays to a top squark and a top quark, $\tilde{g} \to t\tilde{t}_1$. In the models considered, largely motivated by natural supersymmetry, the gluino has a mass around 1-1.5 TeV and the lighter top squark has a mass of a few hundred GeV. Owing to the significant mass gap

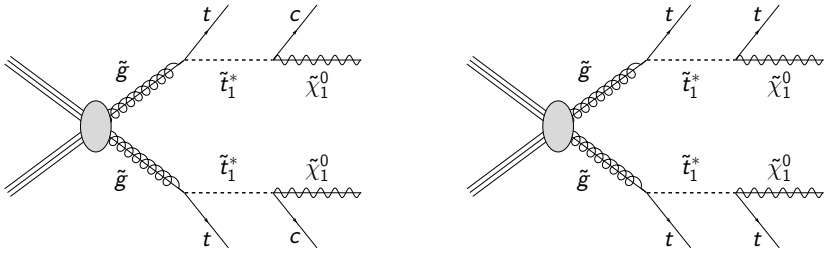

Figure 7.2: Diagram illustrating the T1ttcc (left) and T1t1t (right) simplified models, which address the compressed and stealthy scenarios, respectively.

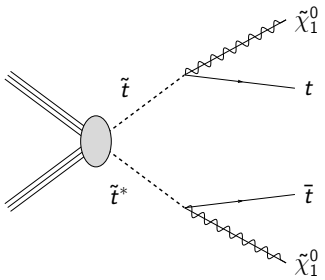

Figure 7.3: Diagram illustrating the T2tt simplified model.

presumed to exist between the gluino and the top squark, the top quark from the gluino to top squark decay will receive a large boost. The top squark then decays to $c\tilde{\chi}_1^0$ for small Δm, or to $t\tilde{\chi}_1^0$ for $\Delta m \approx m_t$. Four-body decays of the top squark are not considered here.

The simplified models (see Section 4.7 for more information) corresponding to the decays $\tilde{g} \to t\tilde{t}_1$, followed by either $\tilde{t}_1 \to c\tilde{\chi}_1^0$ or $\tilde{t}_1 \to t\tilde{\chi}_1^0$, are called *T1ttcc* and *T1t1t*, respectively, and are illustrated in the diagrams in Fig. 7.2. For comparison we show in Fig. 7.3 the diagram for the *T2tt* simplified model, corresponding to direct top squark production in which the top squark decays to $t\tilde{\chi}_1^0$.

As the analysis described in this thesis is the first analysis within CMS to explicitly probe gluino-mediated production of top squark pairs decaying as $\tilde{t}_1 \to c\tilde{\chi}_1^0$, it provides new information about the viability of natural SUSY.

7.2 General strategy

In light of the discussion in Section 7.1, it is expected that boosted top quarks are a promising signature of new physics involving a massive gluino decaying to a relatively light top squark. Boosted objects with high transverse momentum

are characterized by merged decay products separated by $\Delta R \sim 2m/p_{\mathrm{T}}$[1], where m and p_{T} denote the mass and transverse momentum of the mother particle, and ΔR is given in terms of azimuthal angle ϕ and pseudorapidity η as $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. For a separation of $\Delta R = 0.5$, the standard cone size of jets used in CMS, a top quark should thus have a momentum of ∼700 GeV, a value difficult to reach with proton-proton collisions at 8 TeV. Therefore, in order to increase the signal efficiency, we consider instead W bosons from top quark decays, which are required to have a more accessible $p_{\mathrm{T}} \sim 320$ GeV. This results in an increase in the signal efficiency of a factor of about 3–5, depending on the considered mass spectrum. Targeting W bosons instead of top quarks also allows us to use smaller cone sizes, resulting in jets with smaller uncertainties. The p_{T} of the top quark and W boson at generator level without applying any selection, is shown in Figs. 7.4 and 7.5 for several signal models, and the SM $t\bar{t}$ process. We observe that the average p_{T} is higher for the signal than for $t\bar{t}$. Requiring the presence of a boosted W boson will thus be part of the strategy to reduce the SM background. Hadronically decaying boosted W boson candidates will be identified using pruned jet mass [146–148] and a jet substructure observable called N-subjettiness [149]. More details on the W tagging technique will be given in Section 7.4.

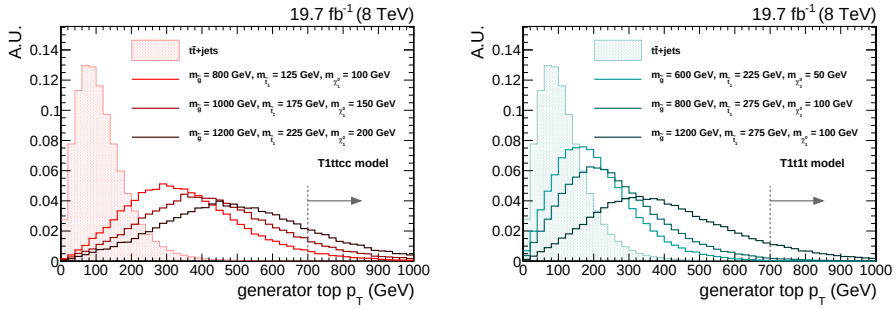

Figure 7.4: Generator level top quark p_{T} for several signal points of the T1ttcc (left) and T1t1t (right) simplified models. The average p_{T} increases as the mass splitting between gluino and top squark increases. The boost of the top quark is larger for the considered signal models in comparison to the SM $t\bar{t}$ background. The range in p_{T} needed to have merged decay products for jets with size $\Delta R = 0.5$ is indicated with an arrow.

The razor kinematic variables $\mathrm{M_R}$ and $\mathrm{R^2}$ (see Section 7.3 for their derivation) are designed to discriminate processes with new heavy particles and missing energy from Standard Model processes. They will be used in this analysis as the main discriminating variables to search for deviations from the SM. We will perform the search in 25 search bins across the high $\mathrm{M_R}$–high $\mathrm{R^2}$ region, using

[1]Considering a heavy object W with mass M that decays to two massless particles a and b, we find $M^2 = 2p_a \cdot p_b = 2E_aE_b(1-\cos\theta_{ab})$. Using small angle approximation this becomes $M^2 = E_aE_b\theta_{ab}^2$. Assigning half of the W energy to both a and b results in $M^2 = \frac{1}{4}E_W^2\theta_{ab}^2$. Translating this relation into the transverse plane, we get $\Delta R = \frac{2M}{p_{\mathrm{T}}^W}$.

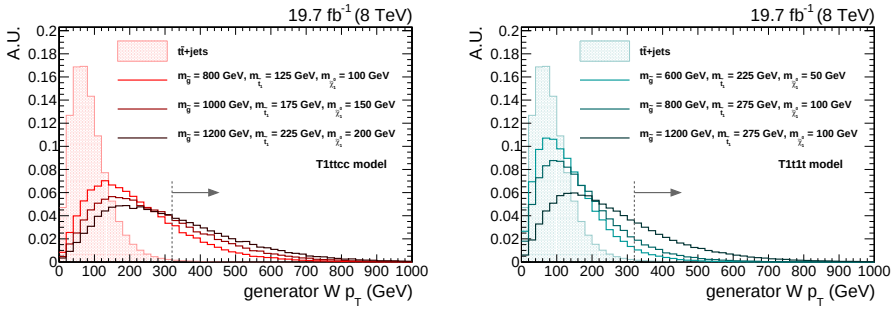

Figure 7.5: Generator level W boson p_{T} for several signal points of the T1ttcc (left) and T1t1t (right) simplified models. The average p_{T} increases as the mass splitting between gluino and top squark increases. The boost of the W boson is larger for the considered signal models in comparison to the SM $t\bar{t}$ background. The range in p_{T} needed to have merged decay products for jets with size $\Delta R = 0.5$ is indicated with an arrow.

hadronic events with at least one boosted W boson and one jet originating from a b quark (i.e. b jet).

Standard model backgrounds in the signal regions are estimated using observations in control regions and global translation factors, calculated from simulated data, that relate the number of events in one region to that in another. Three control regions, Q, W, and T, are defined to select high-purity samples of multijet, $W(\rightarrow \ell\nu)$+jets and $t\bar{t}$ processes, respectively. The background estimation method uses a likelihood-based approach, with a simultaneous sampling of systematic uncertainties which fully takes into account any correlations automatically. The full explanation of the background estimation method will be given in Section 7.7.

7.3 Razor variables

Many extensions of the Standard Model (see Chapter 3) predict the existence of new particles which can be pair-produced in the proton-proton collisions at the LHC. Some of those theories introduce an extra symmetry, such as the R-parity in supersymmetry. A consequence of this symmetry is that the lightest BSM particle must be stable, as it cannot decay to SM particles only. This lightest BSM particle, called LSP in supersymmetric theories, is weakly interacting, and escapes the detector unseen.

This general property leads to a generic class of new physics signatures in which a heavy particle is pair-produced, and decays into visible, i.e. interacting with detector material, SM particles, and an invisible LSP. This signature is illustrated in Fig. 7.6.

A number of kinematical variables targeting this topology have been developed [150–154]. Most of these variables rely on the presence of the invisible LSPs. This causes the visible system to deviate from a dijet topology, resulting

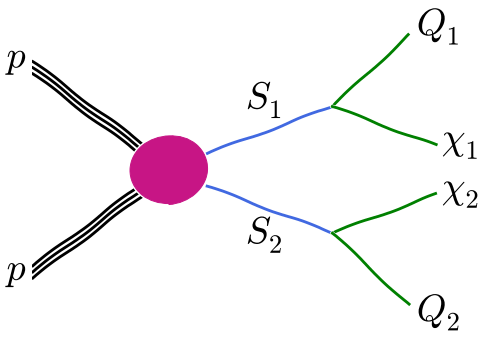

Figure 7.6: Generic new physics signature. Two massive new particles, S_1 and S_2, are produced in pp collisions at the LHC, and consequently decay to a visible system Q_i and an invisible system χ_i.

in possibly large missing transverse momentum, altered angular distributions, etcetera. All of this can be used to distinguish the sought-after signal from the known background processes. Unfortunately, the ultimate goal of reconstructing the masses of the new particles cannot be attained. Because of the escaping LSPs, there is simply not enough information available to fully constrain the problem. What we can do, however, is approximate the mass scale of the involved particles in a collision, which can then be used to check for evidence of new particles. The variables that approximate this mass scale often exhibit a kinematic edge. The *razor variables* [155–158] are no exception in this regard. One advantage the razor variables have over many other variables, is that they also reconstruct the mass scale as a peak, in addition to an edge. In what follows I will derive the two razor variables, denoted $\mathrm{M_R}$ and R^2, which use longitudinal and transverse event information, respectively, to estimate a characteristic mass scale associated with the new particles. The exact use of the razor variables in the razor boost analysis will be discussed in the section about the event selection (Section 7.6.1).

7.3.1 Kinematical configuration and notation

Let us again consider Fig. 7.6. For simplicity, we will assume that the produced particles S_1 and S_2 undergo a two-body decay. Each S_i decays to a visible, Standard Model particle Q_i, and a particle χ_i that escapes the detector. We assume a symmetric decay chain, with the following relations for the masses of the different particles,

$$M_{S_1} = M_{S_2} = M_S \tag{7.1}$$
$$M_{\chi_1} = M_{\chi_2} = M_\chi \tag{7.2}$$
$$M_{Q_1} = M_{Q_2} = 0 \tag{7.3}$$

There are four relevant reference frames for our goal of determining a characteristic mass scale of the new physics process under consideration. The fol-

lowing paragraphs will go through each of these and define the notations that will be used, as well as deriving relations between several variables.

S_1 rest frame From basic two-body decay kinematics it follows that the Q_1 and χ_1 particles are produced back-to-back, with equal magnitude of momentum, in the rest frame of the S_1 particle. This is illustrated in Fig. 7.7.

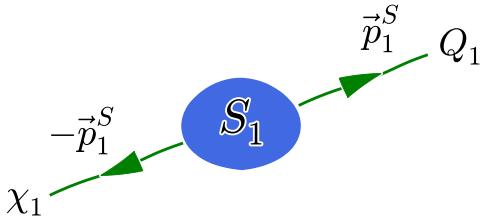

Figure 7.7: Configuration of the S_1 rest frame. The decay products Q_1 and χ_1 are produced back-to-back with momenta $\vec{p}_1^{\,S}$ and $-\vec{p}_1^{\,S}$, respectively.

We can compute the magnitude of this momentum in terms of the new particle masses. To do this we start from the four-vectors P of the Q_1 and χ_1 particles in the S_1 rest frame,

$$P[Q_1] \equiv q_1^S = \{E_{Q_1}^S, \vec{q}_1^{\,S}\}, \tag{7.4}$$

$$P[\chi_1] \equiv \nu_1^S = \{E_{\chi_1}^S, \vec{\nu}_1^{\,S}\}. \tag{7.5}$$

Conservation of energy in the S_1 rest frame leads to

$$E_{Q_1}^S + E_{\chi_1}^S = M_S. \tag{7.6}$$

This can also be expressed as

$$\sqrt{M_{Q_1}^2 + (\vec{q}_1^{\,S})^2} + \sqrt{M_{\chi_1}^2 + (\vec{\nu}_1^{\,S})^2} = M_S. \tag{7.7}$$

Using Eq. 7.2, Eq. 7.3 (massless Q_1), and the equal momenta $|\vec{q}_1^{\,S}| = |\vec{\nu}_1^{\,S}| = |\vec{p}_1^{\,S}|$, the above can be simplified as

$$|\vec{p}_1^{\,S}| = M_S - \sqrt{M_\chi^2 + (\vec{p}_1^{\,S})^2} \tag{7.8}$$

$$(\vec{p}_1^{\,S})^2 = M_S^2 - 2M_S\sqrt{M_\chi^2 + (\vec{p}_1^{\,S})^2} + M_\chi^2 + (\vec{p}_1^{\,S})^2 \tag{7.9}$$

$$2M_S\sqrt{M_\chi^2 + (\vec{p}_1^{\,S})^2} = M_S^2 + M_\chi^2 \tag{7.10}$$

$$4M_S^2(\vec{p}_1^{\,S})^2 = (M_S^2)^2 + 2M_S^2M_\chi^2 + (M_\chi^2)^2 - 4M_S^2M_\chi^2 \tag{7.11}$$

$$(\vec{p}_1^{\,S})^2 = \frac{(M_S^2 - M_\chi^2)^2}{4M_S^2}. \tag{7.12}$$

We thus find for the magnitude of the momentum of Q_1 and χ_1 in the S_1 rest frame

$$|\vec{p}_1^{\,S}| = \frac{M_S^2 - M_\chi^2}{2M_S} \equiv \frac{M_\Delta}{2}, \tag{7.13}$$

where we have defined the characteristic scale M_Δ. This scale is exactly the scale we are interested in. The goal of the **razor variables** is to **express** M_Δ **using lab frame quantities only**. To succeed in this effort, we need to make several, physics-motivated, approximations. These will remove the unknown degrees of freedom, and are further explained in sections 7.3.2 and 7.3.3.

The energy of the Q_1 and χ_1 particles can also be computed easily. From the masslessness of Q_1, we immediately find using Eq. 7.13

$$E^S_{Q_1} = |\vec{p}^S_1| = \frac{M_\Delta}{2}. \tag{7.14}$$

To compute $E^S_{\chi_1}$, we substitute Eq. 7.14 in Eq. 7.6, and find

$$\begin{aligned} E^S_{\chi_1} &= M_S - |\vec{p}^S_1| && (7.15)\\ &= M_S - \frac{M_S^2 - M_\chi^2}{2M_S} && (7.16)\\ &= \frac{2M_S^2 - M_S^2 + M_\chi^2}{2M_S} && (7.17)\\ &= \frac{M_S^2 + M_\chi^2}{2M_S} && (7.18)\\ &= \frac{M_S^2 - M_\chi^2}{2M_S}\frac{M_S^2 + M_\chi^2}{M_S^2 - M_\chi^2}. && (7.19) \end{aligned}$$

We can summarize the four-momenta of Q_1 and χ_1 in the S_1 rest frame as

$$q^S_1 = \frac{M_\Delta}{2}\{1, \vec{u}_1\}, \tag{7.20}$$

$$\nu^S_1 = \frac{M_\Delta}{2}\{R_{S\chi}, -\vec{u}_1\}, \tag{7.21}$$

with

$$R_{S\chi} = \frac{M_S^2 + M_\chi^2}{M_S^2 - M_\chi^2},$$

and $\vec{u}_1$ the unit vector along the Q_1 momentum direction.

S_2 rest frame The discussion of the S_2 rest frame is fully analogous to that of the S_1 rest frame. We again find that

$$q^S_2 = \frac{M_\Delta}{2}\{1, \vec{u}_2\}, \tag{7.22}$$

$$\nu^S_2 = \frac{M_\Delta}{2}\{R_{S\chi}, -\vec{u}_2\}, \tag{7.23}$$

and thus

$$|\vec{p}^S_1| = |\vec{p}^S_2| = \frac{M_\Delta}{2}. \tag{7.24}$$

Centre-of-mass frame In the centre-of-mass (CM) frame of the considered pp collision events the particles S_1 and S_2, which have equal mass (Eq. 7.1), are produced with equal and opposite velocities $\vec{\beta}^{\,\mathrm{CM}}$, as illustrated in Fig. 7.8. The boost $\vec{\beta}^{\,\mathrm{CM}}$ is an indication of how far above threshold the S_i particles are produced, but unfortunately this is an unknown at hadron colliders.

Figure 7.8: Configuration of the centre-of-mass frame. The particles S_1 and S_2 are produced back-to-back with velocities $\vec{\beta}^{\,\mathrm{CM}}$ and $-\vec{\beta}^{\,\mathrm{CM}}$, respectively.

To go from the rest frame of S_1 (S_2) to the CM frame, we need to boost the four-momenta q_1^S and ν_1^S (q_2^S and ν_2^S) to the frame travelling at velocity $\vec{\beta}^{\,\mathrm{CM}}$ ($-\vec{\beta}^{\,\mathrm{CM}}$) with respect to the S_1 (S_2) rest frame. The four-vectors of particles S_1 and S_2 in the centre-of-mass frame are also obtained by boosting according to $\vec{\beta}^{\,\mathrm{CM}}$. They can be written as

$$P[S_1] \equiv s_1^{\mathrm{CM}} = \{E_{S_1}^{\mathrm{CM}}, \vec{s}_{S_1}^{\mathrm{CM}}\} = M_S\,\gamma^{\mathrm{CM}}\,\{1, \vec{\beta}^{\,\mathrm{CM}}\}, \tag{7.25}$$

$$P[S_2] \equiv s_2^{\mathrm{CM}} = \{E_{S_2}^{\mathrm{CM}}, \vec{s}_{S_2}^{\mathrm{CM}}\} = M_S\,\gamma^{\mathrm{CM}}\,\{1, -\vec{\beta}^{\,\mathrm{CM}}\}, \tag{7.26}$$

and satisfy the following

$$(s_1^{\mathrm{CM}} + s_1^{\mathrm{CM}})^2 = \hat{s} = 4(\gamma^{\mathrm{CM}})^2 M_S^2, \tag{7.27}$$

with $\sqrt{\hat{s}}$ the centre-of-mass energy of the collision.

Lab frame The lab frame is the frame where we make our measurements, and is related to the CM frame by a boost $\vec{\beta}^{\,\mathrm{lab}}$. We can decompose this boost into a transverse and longitudinal part as $\vec{\beta}^{\,\mathrm{lab}} = (\vec{\beta}_T, \vec{\beta}_z)$. The four-momenta of the S_i, Q_i and χ_i particles are denoted by s_i^{lab}, q_i^{lab}, and ν_i^{lab} respectively.

7.3.2 Derivation of M_R

As mentioned in the previous section, our goal is to express the characteristic scale M_Δ using lab frame quantities only. Because the problem is kinematically underconstrained, we will need to make some approximations as we work our way from the lab frame to the S_i rest frame, reversing the boosts $\vec{\beta}^{\,\mathrm{lab}}$ and $\vec{\beta}^{\,\mathrm{CM}}$ as we go along.

The models of new physics that we aim to target with the razor variables all predict that the new particles are heavy. This prediction is the basis of the two approximations we will be making.

1. If M_S is large compared to $\sqrt{\hat{s}}$, then the particles S_1 and S_2 will be produced near the $\sqrt{\hat{s}} = 2M_S$ threshold. This means that $\gamma^{\mathrm{CM}} \approx 1$. We will thus assume that $\gamma^{\mathrm{CM}} = 1$, which means that $\vec{\beta}^{\,\mathrm{CM}} \to 0$. The CM frame is thus equal to both S_i rest frames after this first approximation.

2. Neglecting the transverse motion of the colliding partons inside the proton, we can approximate the transverse boost between lab frame and CM frame by

$$|\vec{\beta}_T| \approx \frac{p_{\mathrm{T}}^{\mathrm{ISR}}}{\sqrt{\hat{s}}} \lesssim \frac{p_{\mathrm{T}}^{\mathrm{ISR}}}{2M_S}, \tag{7.28}$$

where $p_{\mathrm{T}}^{\mathrm{ISR}}$ is the magnitude of the vectorial sum of the transverse momentum of the initial state radiation. For large values of M_S we find $\vec{\beta}_T \ll 1$. We thus assume that $\vec{\beta}_T \to 0$, and thus $\vec{\beta}^{\mathrm{lab}} \to \vec{\beta}_z$.

The results of these two approximations is that we only need a longitudinal boost $\vec{\beta}_z$ to take us from the lab frame to the approximate S_i rest frames. In this so-called *rough-approximation frame*, or *R-frame*, we have

$$|\vec{q}_1^{\,R}| = |\vec{q}_2^{\,R}| = \frac{M_\Delta}{2}, \tag{7.29}$$

$$\text{or } E_{Q_1}^R = E_{Q_2}^R = \frac{M_\Delta}{2}, \tag{7.30}$$

cf. Eq. 7.24. Using this constraint, we can compute the boost β^R that will take us from the lab frame to the R-frame. We start from the basic Lorentz transformations for energy and momentum,

$$E_{Q_i}^R = \gamma^R \left(E_{Q_i}^{\mathrm{lab}} - \beta^R \vec{q}_{iz}^{\,\mathrm{lab}}\right), \tag{7.31}$$

$$\vec{q}_{iz}^{\,R} = \gamma^R \left(\vec{q}_{iz}^{\,\mathrm{lab}} - \beta^R E_{Q_i}^{\mathrm{lab}}\right). \tag{7.32}$$

Using Eq. 7.30 we find

$$E_{Q_1}^{\mathrm{lab}} - \beta^R \vec{q}_{1z}^{\,\mathrm{lab}} = E_{Q_2}^{\mathrm{lab}} - \beta^R \vec{q}_{2z}^{\,\mathrm{lab}}, \tag{7.33}$$

and thus for the boost β^R

$$\beta^R = \frac{E_{Q_1}^{\mathrm{lab}} - E_{Q_2}^{\mathrm{lab}}}{\vec{q}_{1z}^{\,\mathrm{lab}} - \vec{q}_{2z}^{\,\mathrm{lab}}}. \tag{7.34}$$

The Lorentz factor γ^R can be expressed as

$$\gamma^R = \frac{1}{\sqrt{1-(\beta^R)^2}} \tag{7.35}$$

$$= \frac{1}{\sqrt{1 - \left(\frac{E_{Q_1}^{\mathrm{lab}} - E_{Q_2}^{\mathrm{lab}}}{\vec{q}_{1z}^{\,\mathrm{lab}} - \vec{q}_{2z}^{\,\mathrm{lab}}}\right)^2}} \tag{7.36}$$

$$= \frac{\vec{q}_{1z}^{\,\mathrm{lab}} - \vec{q}_{2z}^{\,\mathrm{lab}}}{\sqrt{\left(\vec{q}_{1z}^{\,\mathrm{lab}} - \vec{q}_{2z}^{\,\mathrm{lab}}\right)^2 - \left(E_{Q_1}^{\mathrm{lab}} - E_{Q_2}^{\mathrm{lab}}\right)^2}} \tag{7.37}$$

We now define $\mathrm{M_R}$ as

$$\mathrm{M_R} \equiv 2|\vec{q}^{\,R}| = M_\Delta. \tag{7.38}$$

It is interesting to note that M_R is invariant under longitudinal boosts. Using Eq. 7.34 and Eq. 7.37, we can express M_R using lab frame quantities only.

$$
\begin{aligned}
M_R &= E^R_{Q_1} + E^R_{Q_2} && (7.39)\\
&= \gamma^R \left(E^{\rm lab}_{Q_1} + E^{\rm lab}_{Q_2} - \beta^R(\vec{q}^{\,\rm lab}_{1z} + \vec{q}^{\,\rm lab}_{2z})\right) && (7.40)\\
&= \frac{\left(\vec{q}^{\,\rm lab}_{1z} - \vec{q}^{\,\rm lab}_{2z}\right)\left(E^{\rm lab}_{Q_1} + E^{\rm lab}_{Q_2} - \frac{E^{\rm lab}_{Q_1} - E^{\rm lab}_{Q_2}}{\vec{q}^{\,\rm lab}_{1z} - \vec{q}^{\,\rm lab}_{2z}}(\vec{q}^{\,\rm lab}_{1z} + \vec{q}^{\,\rm lab}_{2z})\right)}{\sqrt{\left(\vec{q}^{\,\rm lab}_{1z} - \vec{q}^{\,\rm lab}_{2z}\right)^2 - \left(E^{\rm lab}_{Q_1} - E^{\rm lab}_{Q_2}\right)^2}} && (7.41)\\
&= \frac{\left(\vec{q}^{\,\rm lab}_{1z} - \vec{q}^{\,\rm lab}_{2z}\right)\left(E^{\rm lab}_{Q_1} + E^{\rm lab}_{Q_2}\right) - \left(E^{\rm lab}_{Q_1} - E^{\rm lab}_{Q_2}\right)\left(\vec{q}^{\,\rm lab}_{1z} + \vec{q}^{\,\rm lab}_{2z}\right)}{\sqrt{\left(\vec{q}^{\,\rm lab}_{1z} - \vec{q}^{\,\rm lab}_{2z}\right)^2 - \left(E^{\rm lab}_{Q_1} - E^{\rm lab}_{Q_2}\right)^2}} && (7.42)\\
&= \frac{2\left(\vec{q}^{\,\rm lab}_{1z} E^{\rm lab}_{Q_2} - \vec{q}^{\,\rm lab}_{2z} E^{\rm lab}_{Q_1}\right)}{\sqrt{\left(\vec{q}^{\,\rm lab}_{1z} - \vec{q}^{\,\rm lab}_{2z}\right)^2 - \left(E^{\rm lab}_{Q_1} - E^{\rm lab}_{Q_2}\right)^2}} && (7.43)
\end{aligned}
$$

Our final expression for M_R in lab frame quantities becomes

$$
M_R = 2\sqrt{\frac{\left(\vec{q}^{\,\rm lab}_{1z} E^{\rm lab}_{Q_2} - \vec{q}^{\,\rm lab}_{2z} E^{\rm lab}_{Q_1}\right)^2}{\left(\vec{q}^{\,\rm lab}_{1z} - \vec{q}^{\,\rm lab}_{2z}\right)^2 - \left(E^{\rm lab}_{Q_1} - E^{\rm lab}_{Q_2}\right)^2}}. \qquad (7.44)
$$

As the R-frame is only an approximation of the S_i rest frames, M_R will be distributed around the characteristic scale M_Δ with degrading resolution as the Lorentz factor increases. More generally, the peak value of M_R scales as $\gamma^{\rm CM} M_\Delta$.

We can use M_R as a way to distinguish signal from background, in particular background from QCD multijet production. Let us consider QCD dijet production. In the dijet rest frame we have for the four-momenta of the two jets, k_1 and k_2

$$
k_1 = \frac{\sqrt{\hat{s}}}{2}\{1, \vec{v}\}, \qquad (7.45)
$$

$$
k_2 = \frac{\sqrt{\hat{s}}}{2}\{1, -\vec{v}\}, \qquad (7.46)
$$

where $\sqrt{\hat{s}}$ is the centre-of-mass energy of the partonic subprocess, and $\vec{v}$ is a unit vector along the dijet axis. For this type of event, $M_R = \sqrt{\hat{s}}$, and is thus sharply falling. Signal will thus appear as a peak over a falling background. Given the large cross sections for background processes, and small expected cross sections for signal processes, this discrimination by itself is not sufficient. There is, however, more information available in the event. We have yet to use the transverse degrees of freedom. These will be incorporated in R^2, as explained in the next section.

7.3.3 Derivation of R^2

We will now create a second way to estimate M_Δ, utilizing the transverse information in the event, as encoded in the missing transverse momentum. We start by defining the variable M_{2S} using the four-vectors q_i^{lab} and ν_i^{lab}

$$M_{2S} = \sqrt{\frac{1}{2}\left[(q_1^{\mathrm{lab}} + \nu_1^{\mathrm{lab}})^2 + (q_2^{\mathrm{lab}} + \nu_2^{\mathrm{lab}})^2\right]}. \tag{7.47}$$

As the sum $q_i^{\mathrm{lab}} + \nu_i^{\mathrm{lab}}$ is just the four-vector associated with S_i, we immediately see that $M_{2S} = M_S$. Expanding Eq. 7.47 and using $M_{Q_i} = 0$, we find

$$M_{2S} = \sqrt{\frac{1}{2}\left((q_1^{\mathrm{lab}})^2 + 2q_1^{\mathrm{lab}}\nu_1^{\mathrm{lab}} + (\nu_1^{\mathrm{lab}})^2 + (q_2^{\mathrm{lab}})^2 + 2q_2^{\mathrm{lab}}\nu_2^{\mathrm{lab}} + (\nu_2^{\mathrm{lab}})^2\right)} \tag{7.48}$$

$$= \sqrt{q_1^{\mathrm{lab}}\nu_1^{\mathrm{lab}} + q_2^{\mathrm{lab}}\nu_2^{\mathrm{lab}} + M_\chi^2} \tag{7.49}$$

$$= \sqrt{\begin{aligned} &E_{Q_1}^{\mathrm{lab}} E_{\nu_1}^{\mathrm{lab}} - \vec{q}_{1T}^{\,\mathrm{lab}} \cdot \vec{\nu}_{1T}^{\,\mathrm{lab}} - \vec{q}_{1z}^{\,\mathrm{lab}} \cdot \vec{\nu}_{1z}^{\,\mathrm{lab}} + E_{Q_2}^{\mathrm{lab}} E_{\nu_2}^{\mathrm{lab}} \\ &\qquad - \vec{q}_{2T}^{\,\mathrm{lab}} \cdot \vec{\nu}_{2T}^{\,\mathrm{lab}} - \vec{q}_{2z}^{\,\mathrm{lab}} \cdot \vec{\nu}_{2z}^{\,\mathrm{lab}} + M_\chi^2. \end{aligned}} \tag{7.50}$$

Unfortunately we do not a priori know the mass of the χ_i particles. We will choose $M_\chi = 0$. As a result M_{2S} will now give us a distribution, with endpoint at M_S. For the particular case $M_\chi = 0$, we actually have $M_S = \frac{M_S^2 - M_\chi^2}{M_S} = M_\Delta$. Consequently, the endpoint of M_{2S} gives us a second way to access the characteristic scale M_Δ.

The χ_i particles are assumed to be only weakly interacting. They pass through the detector unseen. When we make the balance of momentum for each event, we can thus infer the existence of these particles. At a hadron collider we can only make this balance in the transverse plane. The transverse component of M_{2S} is given by

$$(M_{2S})_T = \sqrt{|\vec{q}_{1T}^{\,\mathrm{lab}}||\vec{\nu}_{1T}^{\,\mathrm{lab}}| - \vec{q}_{1T}^{\,\mathrm{lab}} \cdot \vec{\nu}_{1T}^{\,\mathrm{lab}} + |\vec{q}_{2T}^{\,\mathrm{lab}}||\vec{\nu}_{2T}^{\,\mathrm{lab}}| - \vec{q}_{2T}^{\,\mathrm{lab}} \cdot \vec{\nu}_{2T}^{\,\mathrm{lab}}}, \tag{7.51}$$

where we have used the assumptions that both Q_i and χ_i are massless. In the considered signal topology we have two unseen particles, χ_1 and χ_2. Experimentally we can only access the sum of their transverse momenta, which in absence of detector effects is given by the missing transverse momentum $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. Making the assumption that the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ is divided equally among both χ_i particles, we find

$$\mathrm{M}_{\mathrm{T}}^{\mathrm{R}} = \sqrt{\frac{|\vec{E}_{\mathrm{T}}^{\mathrm{miss}}|}{2}\left(|\vec{q}_{1T}^{\,\mathrm{lab}}| + |\vec{q}_{2T}^{\,\mathrm{lab}}|\right) - \frac{\vec{E}_{\mathrm{T}}^{\mathrm{miss}}}{2} \cdot \left(\vec{q}_{1T}^{\,\mathrm{lab}} + \vec{q}_{2T}^{\,\mathrm{lab}}\right)}. \tag{7.52}$$

We now define the dimensionless variable R as

$$\mathrm{R} \equiv \frac{\mathrm{M}_{\mathrm{T}}^{\mathrm{R}}}{\mathrm{M}_{\mathrm{R}}}. \tag{7.53}$$

This variable peaks at around 0.5 for signal events, since it is the ratio of two variables that estimate the same scale, with an additional geometric factor to take into account that M_T^R only contains transverse information. For QCD dijet events R equals 0 for an ideal detector. Placing a minimum requirement on this variable can thus be effectively used to suppress background events.

7.3.4 Improved M_R and R^2 definitions

The razor frame as defined in the previous section has a number of useful features, such as M_R being invariant under longitudinal boosts. It also has one important issue which can occur if the approximation $\gamma^{CM} = 1$ breaks down. The issue is visible from the expression of β^R in Eq. 7.34. Looking at that equation, we see that it is possible to get a situation where $|\beta^R| \geq 1$. The boost is then unphysical, and the razor frame ill-defined. We can remedy this issue by not neglecting $\vec{\beta}_T$, the transverse component of $\vec{\beta}^{CM}$.

We again start by making a longitudinal boost β^{L^*} from the lab frame. Then we apply a transverse boost $\vec{\beta}_T^{R^*}$. This boost is applied in opposite directions to the decay products of S_1 and S_2. The two resulting frames are called R*-frames, and have to satisfy the requirement that the magnitude of the momenta of Q_1 and Q_2 in their respective R*-frame are equal.

This constraint can be rewritten as

$$\gamma^{L^*}(E_{Q_1}^{\text{lab}} - E_{Q_2}^{\text{lab}}) - \gamma^{L^*}\beta^{L^*}(q_{1z}^{\text{lab}} - q_{2z}^{\text{lab}}) = \vec{\beta}_T^{R^*} \cdot (\vec{q}_{1T}^{\text{lab}} + \vec{q}_{2T}^{\text{lab}}), \tag{7.54}$$

where we have used the Lorentz transformations for energy corresponding to the two consecutive boosts,

$$E_{Q_i}^{L^*} = \gamma^{L^*}(E_{Q_i}^{\text{lab}} - \beta^{L^*} q_{iz}^{\text{lab}}), \tag{7.55}$$

$$E_{Q_1}^{R^*} = \gamma_T^{R^*}(E_{Q_1}^{L^*} - \vec{\beta}_T^{R^*} \cdot \vec{q}_{1T}^{\text{lab}}), \tag{7.56}$$

$$E_{Q_2}^{R^*} = \gamma_T^{R^*}(E_{Q_2}^{L^*} + \vec{\beta}_T^{R^*} \cdot \vec{q}_{2T}^{\text{lab}}). \tag{7.57}$$

Introducing the unit vector $\hat{\beta}_T^{R^*}$ such that $\vec{\beta}_T^{R^*} = \beta_T^{R^*}\hat{\beta}_T^{R^*}$, we find for the magnitude of the transverse boost

$$\beta_T^{R^*} = \frac{\gamma^{L^*}(E_{Q_1}^{\text{lab}} - E_{Q_2}^{\text{lab}}) - \gamma^{L^*}\beta^{L^*}(q_{1z}^{\text{lab}} - q_{2z}^{\text{lab}})}{\hat{\beta}_T^{R^*} \cdot (\vec{q}_{1T}^{\text{lab}} + \vec{q}_{2T}^{\text{lab}})} \tag{7.58}$$

In analogy with our previous discussion we define the R*-frame mass M_{R^*} as

$$M_{R^*} \equiv 2|\vec{q}_1^{R^*}| = 2|\vec{q}_2^{R^*}| \tag{7.59}$$

$$= \frac{2\gamma^{L^*}\hat{\beta}_T^{R^*} \cdot \left[(E_{Q_1}^{\text{lab}}\vec{q}_{2T}^{\text{lab}} + E_{Q_2}^{\text{lab}}\vec{q}_{1T}^{\text{lab}}) - \beta^{L^*}(q_{1z}^{\text{lab}}\vec{q}_{2T}^{\text{lab}} + q_{2z}^{\text{lab}}\vec{q}_{1T}^{\text{lab}})\right]}{\sqrt{|\hat{\beta}_T^{R^*} \cdot (\vec{q}_{1T}^{\text{lab}} + \vec{q}_{2T}^{\text{lab}})|^2 - (\gamma^{L^*})^2\left[E_{Q_1}^{\text{lab}} - E_{Q_2}^{\text{lab}} - \beta^{L^*}(q_{1z}^{\text{lab}} - q_{2z}^{\text{lab}})\right]^2}} \tag{7.60}$$

To fully compute M_{R^*} we need to pick a value for $\beta_T^{R^*}$ and β^{L^*}. The configurations that led to unphysical M_R values have the property that the momenta

of Q_1 and Q_2 point in the same direction in the transverse plane, with $\vec{\beta}^{\,\mathrm{CM}}$ pointing in the same or the opposite direction. Based on this observation, we choose a direction for $\hat{\beta}_T^{\mathrm{R}^*}$ that maximizes $|\hat{\beta}_T^{R^*} \cdot (\vec{q}_{1T}^{\,\mathrm{lab}} + \vec{q}_{2T}^{\,\mathrm{lab}})|$. The direction that maximizes this quantity is aligned with the direction of $\vec{q}_{1T}^{\,\mathrm{lab}} + \vec{q}_{2T}^{\,\mathrm{lab}}$. Realizing that a unit vector can be expressed as a vector indicating the direction divided by the norm of that vector, we find for $\hat{\beta}_T^{\mathrm{R}^*}$

$$\hat{\beta}_T^{\mathrm{R}^*} = \frac{\vec{q}_{1T}^{\,\mathrm{lab}} + \vec{q}_{2T}^{\,\mathrm{lab}}}{|\vec{q}_{1T}^{\,\mathrm{lab}} + \vec{q}_{2T}^{\,\mathrm{lab}}|} \tag{7.61}$$

We still want $\mathrm{M}_{\mathrm{R}^*}$ to be invariant under longitudinal boosts. Therefore we choose β^{L^*} according to the condition $\frac{\partial \mathrm{M}_{\mathrm{R}^*}}{\partial \beta^{L^*}} = 0$. We find

$$\beta^{L^*} = \frac{q_{1z}^{\mathrm{lab}} + q_{2z}^{\mathrm{lab}}}{E_{Q_1}^{\mathrm{lab}} + E_{Q_2}^{\mathrm{lab}}} \tag{7.62}$$

Substituting Eq. 7.61 and Eq. 7.62 in the expression for $\mathrm{M}_{\mathrm{R}^*}$ we find

$$\mathrm{M}_{\mathrm{R}^*} = \sqrt{\left(E_{Q_1}^{\mathrm{lab}} + E_{Q_2}^{\mathrm{lab}}\right)^2 - \left(q_{1z}^{\mathrm{lab}} + q_{2z}^{\mathrm{lab}}\right)^2 - \frac{(|\vec{q}_{1T}^{\,\mathrm{lab}}|^2 - |\vec{q}_{2T}^{\,\mathrm{lab}}|^2)^2}{|\vec{q}_{1T}^{\,\mathrm{lab}} + \vec{q}_{2T}^{\,\mathrm{lab}}|^2}}. \tag{7.63}$$

We can also express $\gamma_T^{R^*}$ in lab frame observables,

$$\gamma_T^{R^*} = \sqrt{\frac{\left(E_{Q_1}^{\mathrm{lab}} + E_{Q_2}^{\mathrm{lab}}\right)^2 - \left(q_{1z}^{\mathrm{lab}} + q_{2z}^{\mathrm{lab}}\right)^2}{\left(E_{Q_1}^{\mathrm{lab}} + E_{Q_2}^{\mathrm{lab}}\right)^2 - \left(q_{1z}^{\mathrm{lab}} + q_{2z}^{\mathrm{lab}}\right)^2 - \frac{(|\vec{q}_{1T}^{\,\mathrm{lab}}|^2 - |\vec{q}_{2T}^{\,\mathrm{lab}}|^2)^2}{|\vec{q}_{1T}^{\,\mathrm{lab}} + \vec{q}_{2T}^{\,\mathrm{lab}}|^2}}}. \tag{7.64}$$

The peak value of the $\mathrm{M}_{\mathrm{R}^*}$ distribution is at M_Δ, whereas the peak value for $\gamma_T^{R*}\mathrm{M}_{\mathrm{R}*}$ is at $\gamma^{\mathrm{CM}} M_\Delta$, as was the case for M_{R} earlier. This motivates us to redefine M_{R} as

$$\mathrm{M}_{\mathrm{R}} = \gamma_T^{R^*}\mathrm{M}_{\mathrm{R}^*} = \sqrt{\left(E_{Q_1}^{\mathrm{lab}} + E_{Q_2}^{\mathrm{lab}}\right)^2 - \left(q_{1z}^{\mathrm{lab}} + q_{2z}^{\mathrm{lab}}\right)^2}, \tag{7.65}$$

which is still longitudinally invariant, but does not suffer from unphysical boosts. The definition of $\mathrm{M}_{\mathrm{T}}^{\mathrm{R}}$ remains the same, and $\mathrm{R} = \frac{\mathrm{M}_{\mathrm{T}}^{\mathrm{R}}}{\mathrm{M}_{\mathrm{R}}}$ is now defined using the updated M_{R} definition.

7.3.5 Generalization to longer decay chains

In the previous discussion we have always assumed that the new particles decay according to a two-body decay chain. We can of course envision many other scenarios in which this is no longer the case. One could have three- or four-body decays, or long decay chains. All of those cases would result in more than two visible particles in the final state.

In order to use the same razor formalism, we simply need to cluster all the visible particles in the event into two so-called *megajets*. Among all possible

clusterings of particles in two groups, we choose the clustering that minimizes the quadratic sum of invariant masses of the two megajets. This choice was observed to be the most efficient in correctly assigning particles with the same mother to the same megajet. The four-momenta of the megajets, which are used to compute the invariant masses, are defined as the sum of the four-momenta of their constituents. M_R and R^2 are then defined as before, using the megajets as the two visible components.

This approach groups particles travelling in the same direction, and is effective at grouping the decay products of each of the two initially produced particles.

7.4 Boosted W boson tagging

One of the main highlights of the razor boost analysis is the tagging of boosted W bosons in order to access a signal dominated phase space. W bosons either decay to two quarks, or to a lepton and a neutrino. The razor boost analysis is an all-hadronic analysis, which means we do not explicitly consider the leptonic decays. W bosons with low to moderate transverse momentum will thus result in two jets, corresponding to the two clusters of particles resulting from the hadronization of the two quarks. As the p_T of the W boson increases, the separation between the two resulting jets decreases. For high enough momentum, the two jets can no longer be fully resolved with the usual jet definitions, and will be reconstructed as a single jet. This turnover in efficiency between the resolved and merged case is illustrated in Fig. 7.9. Depending on the requirements on the jet multiplicity, losing a jet can result in a loss of signal efficiency. We can, however, also use this effect to our advantage, namely to increase the signal-to-background ratio by requiring the presence of one of these *merged* jets. This, in turn, allows us to relax the jet multiplicity requirements.

The merged jet can be distinguished from other jets by its jet substructure, as illustrated in Fig. 7.10. Jets originating from a W boson should have a two-prong structure, whereas a quark/gluon-initiated jet is not expected to have this structure. In recent years, jet substructure techniques have seen very active developments, and many different algorithms are on the market [160–163]. For the razor boost analysis we will use the CMS recommendation in terms of which techniques to use [159, 164]. We will employ *jet pruning* and a set of variables called *N-subjettiness*. On top of these jet substructure techniques we will also use the jet mass variable to distinguish W boson-initiated jets from quark/gluon-initiated jets. The following subsections will go through the different parts of the W tagging definition, providing a more detailed explanation for each.

7.4.1 Jet algorithm

In order to identify boosted W bosons, we will use a different jet clustering algorithm than what is used for the standard jet definition (see Section 6.4.2). Jets will be clustered with FastJet 3.0.1. [114], from the PF candidates,

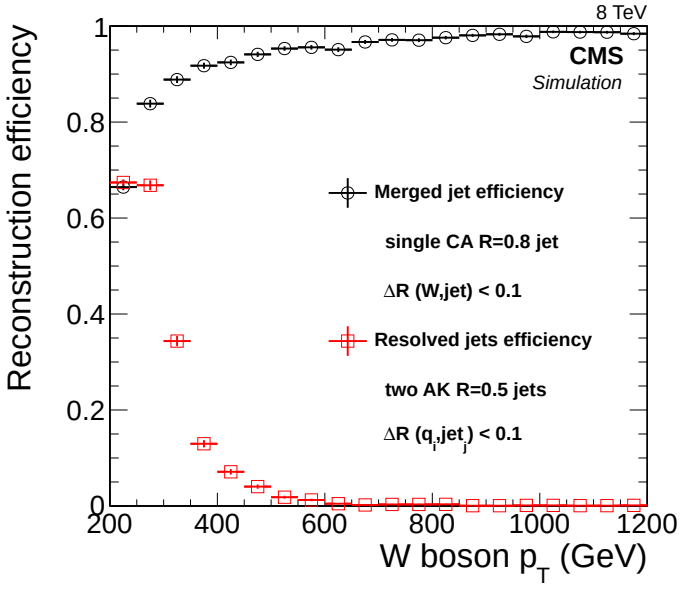

Figure 7.9: Efficiency to reconstruct a CA8 jet within $\Delta R < 0.1$ of a generated W boson, and the efficiency to reconstruct two AK5 jets within $\Delta R < 0.1$ of the generated quarks from longitudinally polarized W bosons, as a function of the p_{T} of the W boson [159]. The loss in efficiency for the resolved case is clearly visible for high p_{T} W bosons.

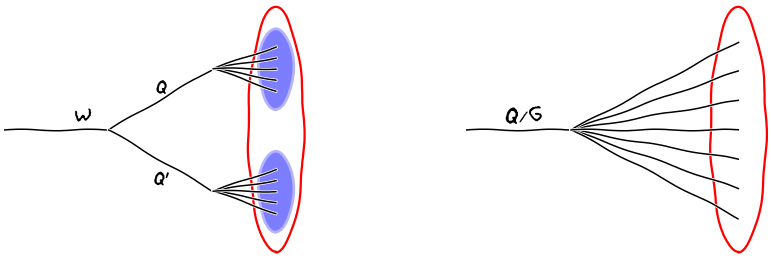

Figure 7.10: The jet substructure of a W-initiated jet differs from a quark/gluon-initiated jet.

using the Cambridge-Aachen (CA) algorithm [165] with a size parameter of 0.8. Henceforth, we will call these jets *CA8 jets*.

Cambridge-Aachen jet algorithm. The Cambridge-Aachen jet algorithm is a sequential recombination algorithm that uses the distance measure d_{ij} between two constituents i and j,

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2}, \tag{7.66}$$

with R the size parameter of the resulting jets, and

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2, \tag{7.67}$$

where y, ϕ are the rapidity (defined in Eq. 6.27) and azimuthal angle. The distance between constituent i and the beam is given by $d_{iB} = 1$. As is clear from the above, these distance measures only use angular information, unlike for the k_T and anti-k_T algorithms, which use a p_T-weighted distance.

The jet algorithm starts by computing the minimum distance d_{ij}, across all i, j. If $\min d_{ij} < d_{iB}$, then we combine constituents i and j into a new constituent whose four-momentum is the sum of the four-momenta of i and j, and repeat the process. Otherwise, we call i a jet and move it from the list of constituents to be clustered to the list of final jets. The process is repeated with the remaining constituents, until none remain.

Jet energy corrections for these CA8 jets are derived from the standard anti-k_T jets with size parameter $R = 0.7$. Simulations show that the corrections are valid for CA8 jets and have an additional uncertainty no greater than 2% [166, 167].

7.4.2 Jet pruning

Jet pruning [146, 147] is a particular kind of jet grooming. Jet grooming techniques are designed to reduce the impact of contributions from the underlying event (UE), pileup (PU), and low-p_T gluon radiation. These kinds of contributions to jets are typically soft and diffuse, and increase the jet energy proportional to the jet area. Grooming techniques reduce the jet area without affecting the core components. This means that the resulting jets are less sensitive to these soft contributions, but still reflect the kinematics of the original, hard process.

During jet pruning the constituents of the jet are reclustered with the CA algorithm, using the same distance parameter as used for the original jets (here $R = 0.8$), but with additional conditions beyond those of the standard algorithm. In particular, the softer and larger-angle of the two particles i and j to be merged is removed when the following conditions are satisfied:

$$z_{ij} = \frac{\min(p_\mathrm{T}^i, p_\mathrm{T}^j)}{p_\mathrm{T}^i + p_\mathrm{T}^j} < z_\mathrm{cut}, \tag{7.68}$$

$$\Delta R_{ij} > D_\mathrm{cut} \equiv \alpha \frac{m_J}{p_\mathrm{T}}, \tag{7.69}$$

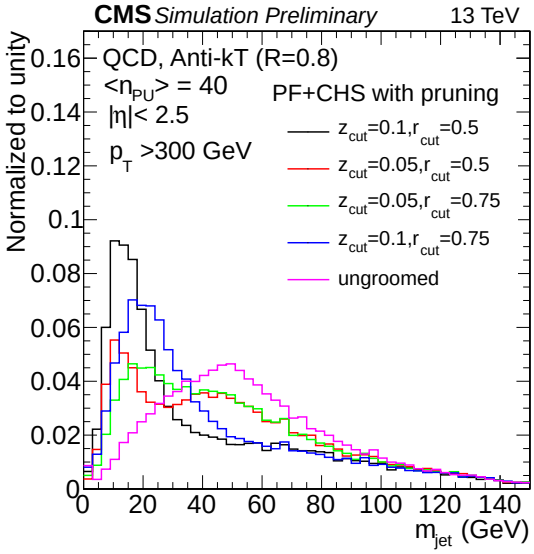

Figure 7.11: Jet mass distribution of QCD jets with $p_{\mathrm{T}}(\mathrm{gen}) > 300\,\mathrm{GeV}$ for jet pruning with different parameters, starting from PF jets with charged-hadron subtraction applied. The fully ungroomed mass distribution is also shown for comparison. Here, the jets were clustered using the anti-k_{T} algorithm, but the same picture holds for Cambridge-Aachen jets. Figure taken from Ref. [111].

where m_J and p_{T} are the original mass and transverse momentum of the reclustered jet, ΔR_{ij} is defined as in Eq. 7.67, and z_{cut} and α are parameters of the algorithm, chosen to be 0.1 and 0.5, respectively [148].

The resulting pruned jet is used as further input to our W boson tagger. For the W decay products to be collimated, we need a large transverse momentum. We will therefore require that the pruned jets have $p_{\mathrm{T}} > 200\,\mathrm{GeV}$. Because of the reduction of the effect of UE and PU, the jet mass variable as computed from the constituents of the jet after jet pruning has a much better behaviour than if it was computed from the unpruned jets, as seen on Fig. 7.11. Jet pruning shifts the jet mass of QCD jets to smaller values, while maintaining the jet mass for W jets close to the W boson mass.

We will make the requirement that the pruned jet mass is consistent with the W boson mass,

$$70 < m_{\text{pruned jet}} < 100\,\mathrm{GeV}. \tag{7.70}$$

Here, we have deviated from the standard interval used in CMS, starting at 60 GeV, as we found that for our kinematical region and signal topology we achieve better signal to background discrimination when increasing the lower cut value to 70 GeV.

7.4.3 N-subjettiness

Requiring the jet mass to be consistent with the W boson mass already results in a good discrimination between W boson and quark/gluon-initiated jets. We can, however, still do better. A boosted QCD jet with a mass around 80 GeV usually originates from a single hard parton and acquires mass through large-angle soft splittings. The energy pattern for this process will differ from the two-prong pattern that is found in boosted W jets. The set of N-subjettiness observables τ_N [149] aims to exploit this difference in expected energy flow to differentiate between W boson and quark/gluon-initiated jets by counting the number of hard lobes of energy within a jet.

N-subjettiness is computed under the assumption that the jet has N subjets, and is the p_{T}-weighted ΔR distance between each jet constituent and its nearest subjet axis:

$$\tau_N = \frac{1}{R_0 \sum_k p_{T,k}} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, ...\Delta R_{N,k}), \tag{7.71}$$

where R_0 is the original jet distance parameter (0.8 in our case) and k runs over all constituent particles of the jet. The subjet axes are obtained by running the exclusive k_T algorithm [168, 169] using FastJet. The exclusive k_T algorithm differs from the inclusive version in two ways: if at a given clustering step $d_{iB} < \min_j d_{ij}$, then constituent i is discarded, rather than added to the jet collection; and the clustering stops when the desired number of jets (N) is reached. The resulting axes can be further optimized to minimize the N-subjettiness value. In accordance to the CMS recommendation, we use a one-pass optimization of the exclusive k_T axes [170].

The variables τ_N quantify the consistency of the jet having N or fewer subjets. They have a small value (close to 0) if the original jet is consistent with having N or fewer subjets, because almost every jet constituent will be close in ΔR to its own true subjet. As we are interested in discriminating boosted W bosons, with two subjets, from quark/gluon jets, which have a single subjet, we will use the variables τ_2 and τ_1, as obtained from the unpruned CA8 jets. It has been shown that the ratio of the τ_N variables are better discriminators than the separate variables [149]. We will thus require that the ratio τ_2/τ_1 is small. To ensure that the N-subjettiness ratio as computed from the unpruned jet collection is assigned to the correct pruned jet, we find the highest p_{T} unpruned jet that is within $\Delta R = 0.7$ of the considered pruned jet. The τ_2/τ_1 distribution for highly boosted and longitudinally polarized W bosons and for inclusive QCD jets is shown in Fig. 7.12.

7.4.4 W boson tagging definitions

In the razor boost analysis we will employ a boosted W boson tagger, utilizing the techniques outlined in the previous sections, to identify events that are consistent with the presence of a high p_{T}, hadronically decaying W boson. A given pruned CA8 jet is W tagged if it has $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, $|\eta| < 2.4$, $70 < m_{\mathrm{jet}} < 100\,\mathrm{GeV}$, and the corresponding unpruned jet satisfies $\tau_2/\tau_1 < 0.5$. This

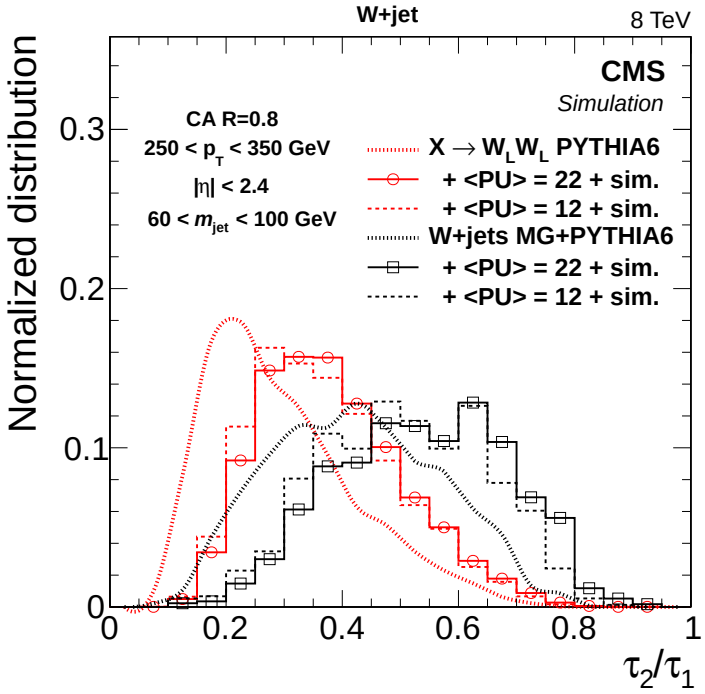

Figure 7.12: Distributions of the N-subjettiness ratio τ_2/τ_1 in simulated samples of highly boosted and longitudinally polarized W bosons and inclusive QCD jets expected in the W+jet topology (i.e. with leptonically decaying W bosons recoiling off a hard jet). The distribution is shown after a selection on the pruned jet mass of $60 < m_{\text{pruned jet}} < 100\,\text{GeV}$. Note that this is slightly different from what is applied in the razor boost analysis. Thick dashed lines represent the generator predictions without pileup interactions and without CMS detector simulation. The histograms are the expected distributions after full CMS simulation with pileup corresponding to an average number of 12 and 22 interactions [159].

definition is the same as was used previously in a search for massive resonances in dijet systems containing jets tagged as a W or Z boson [171,172]. The precise definition of this W boson tagger is summarized in Table 7.1.

As explained in Section 7.2, we will use three control regions to select data samples enriched in QCD multijet, $t\bar{t}$, and $W(\to l\nu)$+jets, in order to help model the SM backgrounds. QCD multijet and leptonically decaying W+jets events are not expected to have jets with a two-prong substructure. Therefore, our W boson tagging definition will not be very efficient in selecting these processes. To remedy this, we slightly modify our W tagger. We define W boson *anti-tagged* jets (aW) by taking the complement of the τ_2/τ_1 requirement, and define W boson *mass-tagged* jets (mW) by dropping that requirement all together. These definitions allow a more efficient selection of background

processes, while remaining in a similar kinematic regime. How these taggers will be used exactly will be explained in Section 7.6.4 when discussing the event selection. A summary of their definitions can be found in Table 7.1 as well.

Table 7.1: Boosted W tagging definitions. The input jet collection is either the pruned or unpruned CA8 jet collection with charged-hadron subtraction applied.

	W	aW	mW
Pruned	$p_{\mathrm{T}} > 200$ $\lvert\eta\rvert < 2.4$ $70 < m_{\mathrm{jet}} < 100$	$p_{\mathrm{T}} > 200$ $\lvert\eta\rvert < 2.4$ $70 < m_{\mathrm{jet}} < 100$	$p_{\mathrm{T}} > 200$ $\lvert\eta\rvert < 2.4$ $70 < m_{\mathrm{jet}} < 100$
Unpruned	$\tau_2/\tau_1 < 0.5$	$\tau_2/\tau_1 \geq 0.5$	-

7.4.5 W boson tagging scale factors

It has been observed by previous CMS analyses that the W boson tagging efficiency is not the same in data and in simulation. The distributions that are at the root of this disagreement are shown on Fig. 7.13. To account for the discrepancies, we need to derive data/MC scale factors and associated uncertainties corresponding to each of the W boson tagging, mass-tagging and anti-tagging definitions listed in Table 7.1. These scale factors are not process-independent. They will be different for processes that include hadronically decaying W bosons, such as $t\bar{t}$ or the signal, compared to processes which do not have W bosons in their final state, such as QCD multijet production. For processes without real hadronically decaying W bosons, any tagged jet is necessarily a misidentified, or *fake*, W boson tag. For those processes we will speak of the W boson tagging fake rate scale factors, where the fake rate is defined as the probability to tag, with one of the used W tagging definitions, a jet not coming from a hadronically decaying W boson. One last consideration concerns the signal simulation. As the signal is simulated with FastSim, we need an additional scale factor to correct for differences in the modelling of the W tagger between FastSim and FullSim. In the following subsections every scale factor will be listed in more detail, including how it was derived and how it will be used in the analysis.

W boson tag efficiency scale factor

The W boson tag efficiency scale factor will be used to correct processes with real hadronically decaying W bosons. For the backgrounds this is mainly for the $t\bar{t}$ process, but also single top and $t\bar{t}$ in association with a W or Z boson are considered. This scale factor is of course also used for the signal processes.

The W boson tag efficiency scale factor is only applied to the simulation in the S and T region, see Sections 7.6.3 and 7.6.4, as those are the regions that utilize the W tagging definition in their selection criteria. It is also only

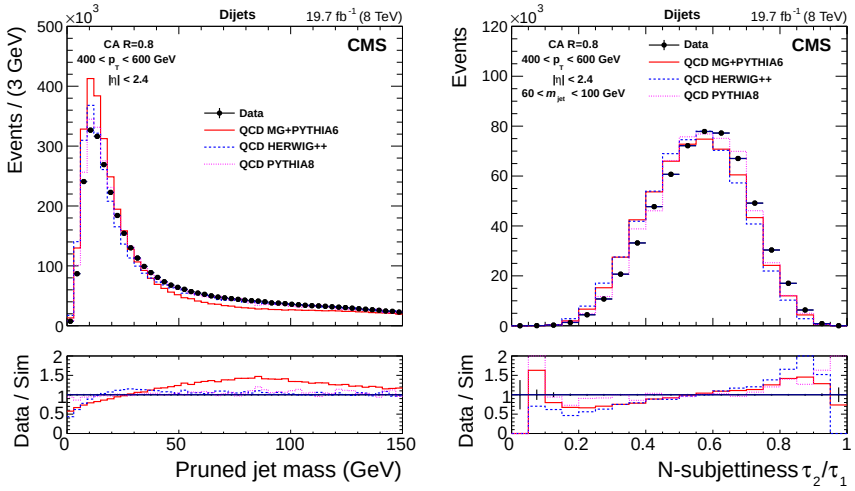

Figure 7.13: Pruned jet mass (left) and N-subjettiness ratio τ_2/τ_1 (right) distributions in data and simulation for dijet events. MG denotes the MadGraph generator, and is the option used for the razor boost analysis. The relative deviations between data and simulation are plotted at the bottom of each figure [159].

applied to events for which the W boson tagged jet is matched (within a cone of $\Delta R = 0.8$) to a generator level hadronically decaying W boson. In case no match was found, we apply the W boson tag fake rate scale factor.

As we use the same W boson tagging definition as was used in a search for massive resonances in dijet systems containing W tagged jets [171], we can directly apply the scale factor that was derived for that study. The method used to obtain the scale factor is outlined in Ref. [164]. The W boson tag efficiency scale factor SF_{Wtag} is given by

$$SF_{\mathrm{Wtag}} = 0.86 \pm 0.07. \tag{7.72}$$

W boson tag efficiency FullSim/FastSim scale factor

For our signal samples, which are produced with FastSim, we have derived an additional W tag efficiency FullSim/FastSim scale factor, $SF_{\mathrm{Full/Fast}}$, which depends on the p_{T} of the CA8 jet. This scale factor corrects for the different modelling of jets, jet substructure, etcetera, in FastSim with respect to FullSim. The product of SF_{Wtag} and $SF_{\mathrm{Full/Fast}}$ will be applied to the signal simulation.

To compute the W boson tag efficiency FullSim/FastSim scale factor we use a sample of $t\bar{t}$ events simulated with both FullSim and FastSim. A FastSim versus FullSim comparison of the distributions of the pruned jet mass, the N-subjettiness variables τ_1, τ_2 and their ratio τ_2/τ_1, both before and after requiring the jets to satisfy the pruned jet mass window, is shown in Figs. 7.14–7.17. It is clear that the agreement is not perfect. The τ_2/τ_1 distribution for

FastSim is shifted with respect to FullSim. This disagreement will then of course be translated into the efficiencies, and thus the need for a scale factor arises.

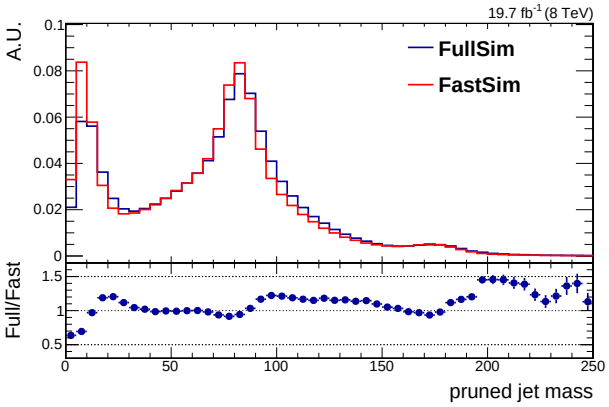

Figure 7.14: Pruned jet mass distribution for FastSim and FullSim $t\bar{t}$.

These figures also illustrate some of the features of the N-subjettiness variables. Consider first the τ_1 distribution in Fig. 7.15. Without a jet mass requirement, the distribution is quite broad and bimodal. Once the jet mass is required to be consistent with the W boson mass, the lower part of the distribution disappears. This illustrates that quark/gluon jets are expected to have only a single subjet, resulting in a small τ_1 value. For jets that result from the decay of a W boson, τ_1 takes on a larger value. This effect is not seen for the τ_2 distributions (Fig. 7.16), which is expected because τ_2 quantifies the compatibility with having two or fewer subjets. The ratio τ_2/τ_1, shown on Fig. 7.17, also displays the expected behaviour: the part of the distribution at high values is removed when requiring the jet mass to be within the W mass window, and thus when selecting more jets with two-prong decays.

The procedure to determine the W boson tagging efficiency for both FastSim and FullSim is the following:

1. Filter the events at the generator level, requiring the presence of exactly one hadronically decaying W boson.

2. For the generated W boson, find the closest reconstructed CA8 jet, and require that it be within $\Delta R = 0.8$ from the W boson. If no such jet exists, the event is discarded.

3. Require that there be no (generator-level) b quark from the top quark decay within the cone of the selected CA8 jet. (We wish to select boosted W bosons only, not boosted top quarks.)

4. For the events that pass the above selection, consider the p_{T} distribution of the CA8 jet at two selection levels:

   - no additional selection

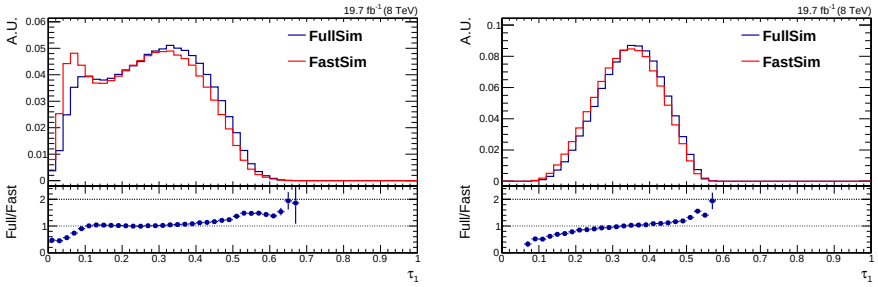

Figure 7.15: Distribution of τ_1 before (left) and after (right) requiring the pruned CA8 jet to lie within the W mass window, $70 < m_{\text{jet}} < 100\,\text{GeV}$, for FastSim and FullSim $t\bar{t}$.

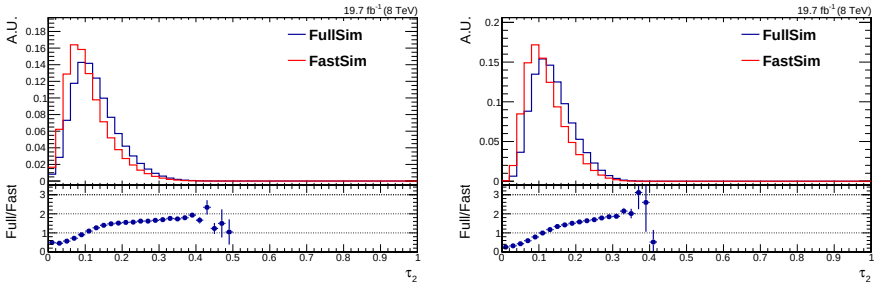

Figure 7.16: Distribution of τ_2 before (left) and after (right) requiring the pruned CA8 jet to lie within the W mass window, $70 < m_{\text{jet}} < 100\,\text{GeV}$, for FastSim and FullSim $t\bar{t}$.

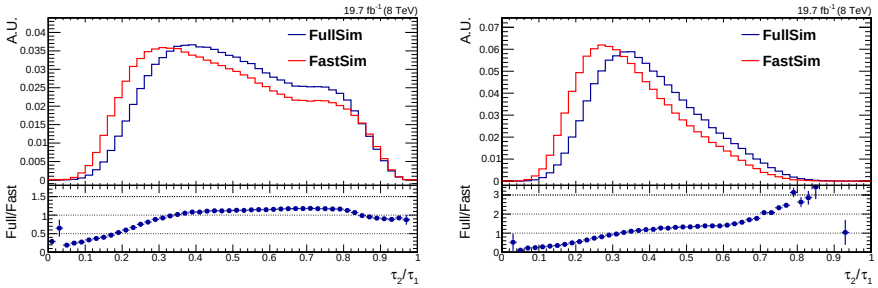

Figure 7.17: Distribution of τ_2/τ_1 before (left) and after (right) requiring the pruned CA8 jet to lie within the W mass window, $70 < m_{\text{jet}} < 100\,\text{GeV}$, for FastSim and FullSim $t\bar{t}$.

- $70 < m_{\text{jet}} < 100\,\text{GeV}$ and $\tau_2/\tau_1 < 0.5$

5. By dividing those p_{T} distributions we obtain the W boson tagging efficiency.

To derive the FullSim/FastSim scale factor for the W boson tagging efficiency, we divide the efficiencies ϵ obtained in FullSim and FastSim:

$$SF_{\text{Full/Fast}}(p_{\text{T}}) = \frac{\epsilon_{\text{FullSim}}(p_{\text{T}})}{\epsilon_{\text{FastSim}}(p_{\text{T}})}. \tag{7.73}$$

A graphical representation of the W boson tag efficiency in FastSim and FullSim is shown on Fig. 7.18 for a fine and more coarse binning in CA8 jet p_{T}. The resulting scale factor is shown for the final, coarse binning that will be used to rescale the signal simulation. Table 7.2 summarizes the W boson tag efficiency FullSim/FastSim scale factor with its statistical uncertainty.

Table 7.2: Summary of FullSim/FastSim scale factor for the W tag efficiency.

CA8 jet p_{T} (GeV)	$SF_{\text{Full/Fast}}$
$[200 - 250[$	0.952 ± 0.010
$[250 - 350[$	0.912 ± 0.012
$[350 - ...]$	0.891 ± 0.026

W boson tag fake rate scale factor

The W boson tag fake rate scale factor is meant to correct processes that do not have hadronically decaying W bosons in their final state. As the fake rate depends on the composition of the sample, this scale factor has to be derived for each analysis separately. We will thus need to obtain a sample of events containing misidentified W boson jets to derive a dedicated scale factor for the razor boost analysis. A multijet-enriched control region is defined, using the following selection:

- no loose leptons,
- no b tagged (CSVL) jets,
- at least 3 AK5 jets,
- at least one AK5 jet with $p_{\text{T}} > 200\,\text{GeV}$,
- small minimum azimuthal angle between the $\vec{E}_{\text{T}}^{\text{miss}}$ and the leading three jets,
  $\Delta\phi_{min} < 0.3$.

This selection is similar to the baseline selection employed in the rest of the analysis. The kinematic regime, and the composition of the sample will thus also be similar. The main difference is that we have not applied any selection on the razor variables M_{R} or R^2, in order to retain a higher statistical power.

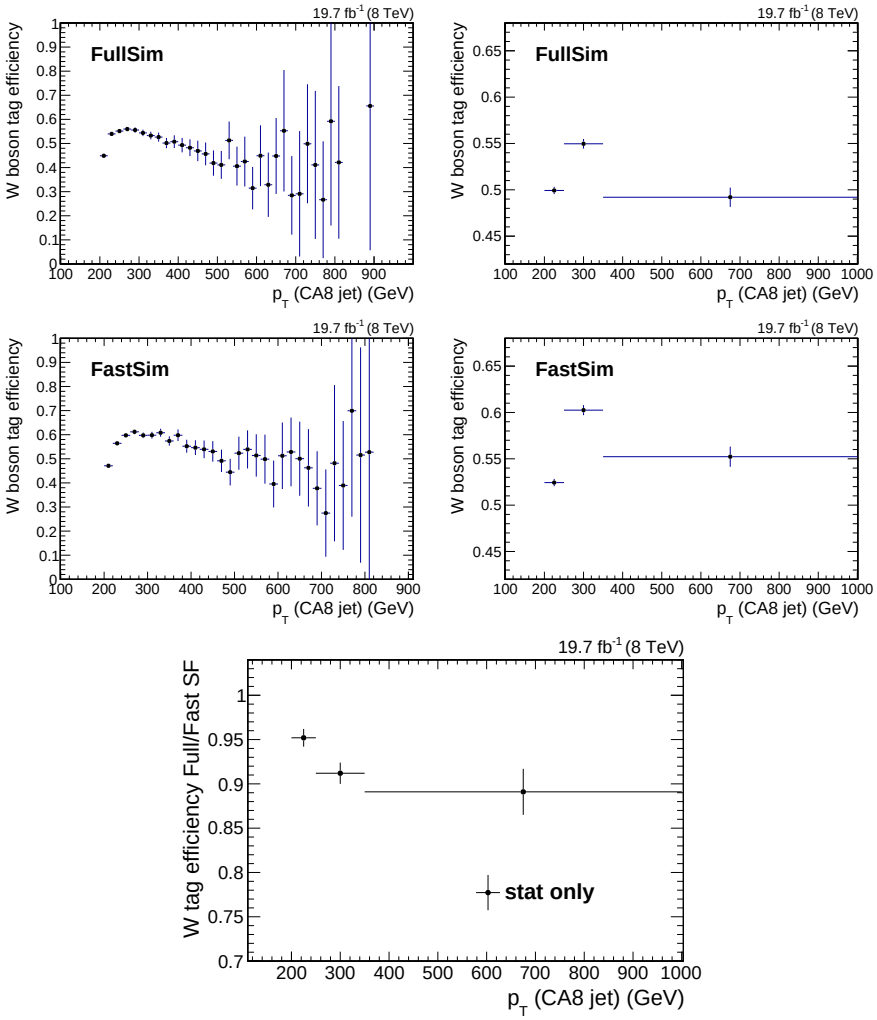

Figure 7.18: [top] W boson tag efficiency versus CA8 jet p_{T}, with two different binnings, obtained from FullSim $t\bar{t}$ events as described in the text. The shown uncertainties are statistical only. [middle] W boson tag efficiency versus CA8 jet p_{T}, with two different binnings, obtained from FastSim $t\bar{t}$ events as described in the text. The shown uncertainties are statistical only. [bottom] W boson tag FullSim/FastSim efficiency scale factor versus CA8 jet p_{T}. The shown uncertainties are statistical only.

To obtain the fake rates ϵ for W boson tagging we use the leading CA8 jet in each event, and check whether it is tagged by the W boson tagger. After obtaining the fake rates in both data and simulation, we compute the scale factor as their ratio,

$$SF^{\text{fake}}_{\text{Wtag}}(p_{\text{T}}) = \frac{\epsilon^{\text{data}}(p_{\text{T}})}{\epsilon^{\text{simulation}}(p_{\text{T}})}. \tag{7.74}$$

In the calculation of the uncertainties on this scale factor we include the statistical uncertainty, as well as the trigger efficiency and jet energy scale uncertainties for both AK5 and CA8 jets. All three uncertainties are varied up (down) at the same time to get the overall up (down) systematic uncertainty. The fake rate in data and simulation, as well as the resulting W boson tag fake rate scale factor are shown in Fig. 7.19. As we can see from the figure, there is a drop in the scale factor just above a p_{T} of 300 GeV. This is a result of a residual mismodelling of the trigger efficiency.

The W boson tag fake rate scale factor is applied in the S and T region, as is the case for the W boson tag efficiency scale factor. It is applied to all simulated samples without real hadronically decaying W bosons, such as multijet production, $W(\rightarrow \ell\nu)$+jets, $\text{Z}/\gamma^*(\rightarrow \ell\ell)$+jets, etcetera.

W boson mass-tag fake rate scale factor

This scale factor corresponds to the W boson mass-tagging definition, which will be used in the W control region. As we will see further, the W region is dominated by leptonically decaying W bosons. The mass-tagged jet, therefore, originates from a quark or gluon jet, and is a misidentified W boson jet. For this reason we will derive the scale factor for the W boson mass-tag fake rate from the same multijet-enriched region as was used to derive the scale factor for the W boson tag fake rate. The same method as before is applied, the only difference being the use of the W boson mass-tag definition instead of the W boson tag definition.

The resulting scale factor, $SF^{\text{fake}}_{\text{Wmasstag}}$, is again a function of the CA8 jet p_{T}, and will be applied to all simulated samples in the W region. Figure 7.20 shows the W boson mass-tag fake rate in data and simulation, and the corresponding scale factor versus CA8 jet p_{T}. The scale factor with associated statistical and systematic uncertainties is listed in Table 7.3. The systematic uncertainty includes the trigger efficiency uncertainty and the uncertainty on the jet energy scale corrections. The drop in the scale factor just above a p_{T} of 300 GeV is also visible here.

W boson anti-tag fake rate scale factor

This scale factor corresponds to the W boson anti-tagging definition, which will be used in the Q control region. This region is dominated by QCD multijet production. Consequently, the W anti-tagged jet originates from a quark/gluon jet and is a misidentified W boson jet. Once more, we use the same procedure and the multijet-enriched region as for the W boson tag fake rate scale factor.

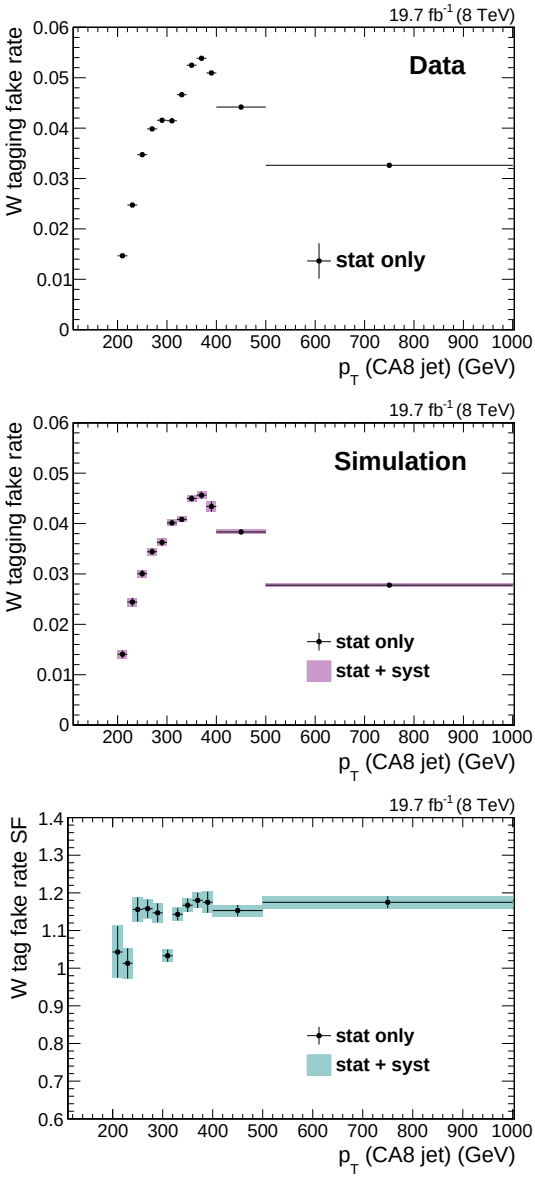

Figure 7.19: [top] Misidentification probability according to W tagging for a CA8 jet versus jet p_{T} obtained from a multijet-enriched control region in data as described in the text. The shown uncertainties are statistical only. [middle] W boson tag fake rate obtained from simulation. The uncertainty band includes statistical and systematic uncertainties. [bottom] Scale factor for W tag fake rate versus CA8 jet p_{T} obtained from a multijet-enriched control region as described in the text. The uncertainty band includes statistical and systematic uncertainties.

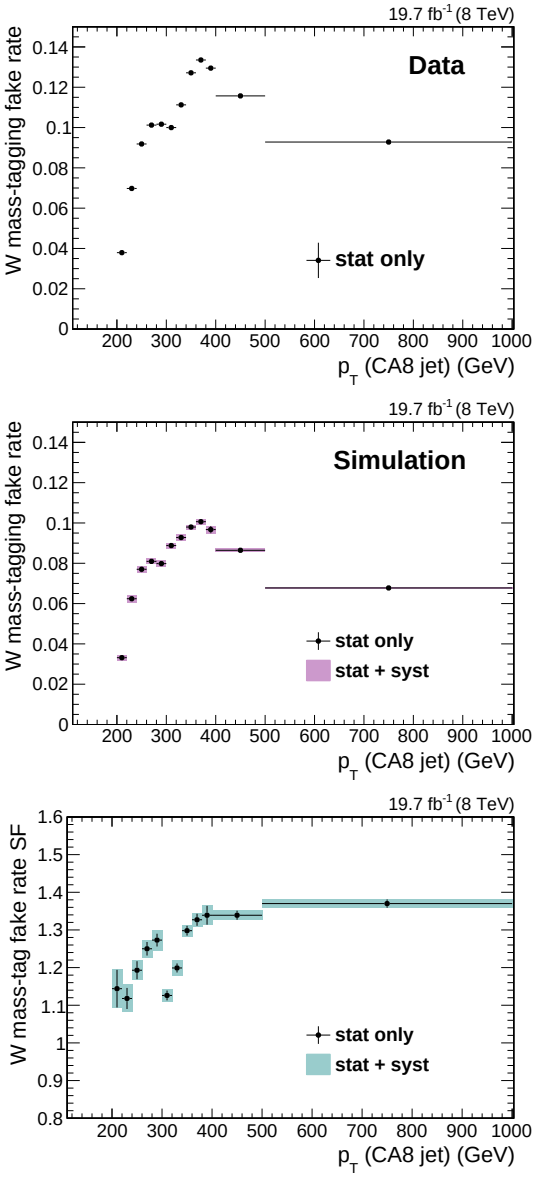

Figure 7.20: [top] Misidentification probability according to W mass-tagging for a CA8 jet versus jet p_{T} obtained from a multijet-enriched control region in data as described in the text. The shown uncertainties are statistical only. [middle] W boson mass-tag fake rate according to simulation. The uncertainty band includes statistical and systematic uncertainties. [bottom] Scale factor for W mass-tag fake rate versus CA8 jet p_{T} obtained from a multijet-enriched control region as described in the text. The uncertainty band includes statistical and systematic uncertainties.

Table 7.3: W boson mass-tag fake rate scale factor, binned in p_{T}. The breakdown in statistical and systematic uncertainties is shown.

CA8 jet p_{T} (GeV)	$SF^{\mathrm{fake}}_{W\mathrm{masstag}}$
$[200-220[$	$1.144 \pm 0.050\,(\mathrm{stat}) \pm 0.012\,(\mathrm{sys})$
$[220-240[$	$1.118 \pm 0.028\,(\mathrm{stat}) \pm 0.024\,(\mathrm{sys})$
$[240-260[$	$1.193 \pm 0.024\,(\mathrm{stat}) \pm 0.008\,(\mathrm{sys})$
$[260-280[$	$1.250 \pm 0.018\,(\mathrm{stat}) \pm 0.015\,(\mathrm{sys})$
$[280-300[$	$1.273 \pm 0.017\,(\mathrm{stat}) \pm 0.021\,(\mathrm{sys})$
$[300-320[$	$1.126 \pm 0.013\,(\mathrm{stat}) \pm 0.010\,(\mathrm{sys})$
$[320-340[$	$1.199 \pm 0.012\,(\mathrm{stat}) \pm 0.017\,(\mathrm{sys})$
$[340-360[$	$1.298 \pm 0.013\,(\mathrm{stat}) \pm 0.007\,(\mathrm{sys})$
$[360-380[$	$1.327 \pm 0.016\,(\mathrm{stat}) \pm 0.008\,(\mathrm{sys})$
$[380-400[$	$1.339 \pm 0.025\,(\mathrm{stat}) \pm 0.007\,(\mathrm{sys})$
$[400-500[$	$1.339 \pm 0.012\,(\mathrm{stat}) \pm 0.005\,(\mathrm{sys})$
$[500-...]$	$1.370 \pm 0.011\,(\mathrm{stat}) \pm 0.001\,(\mathrm{sys})$

The W boson anti-tag fake rate scale factor, $SF^{\mathrm{fake}}_{\mathrm{Wantitag}}$, will be applied to all simulated samples in the Q region. Figure 7.21 shows the W boson anti-tag fake rate in data and simulation, and the corresponding scale factor versus CA8 jet p_{T}. As before, we observe a drop in the scale factor just above a p_{T} of 300 GeV. The scale factor, and breakdown of statistical and systematic uncertainties, is listed in Table 7.4. The included systematic uncertainties are those stemming from the trigger efficiency and jet energy scale corrections.

Table 7.4: W boson anti-tag fake rate scale factor, binned in p_{T}. The uncertainties are broken down in their statistical and systematic component.

CA8 jet p_{T} (GeV)	$SF^{\mathrm{fake}}_{W\mathrm{antitag}}$
$[200-220[$	$1.217 \pm 0.072\,(\mathrm{stat}) \pm 0.032\,(\mathrm{sys})$
$[220-240[$	$1.186 \pm 0.037\,(\mathrm{stat}) \pm 0.046\,(\mathrm{sys})$
$[240-260[$	$1.216 \pm 0.033\,(\mathrm{stat}) \pm 0.011\,(\mathrm{sys})$
$[260-280[$	$1.319 \pm 0.024\,(\mathrm{stat}) \pm 0.019\,(\mathrm{sys})$
$[280-300[$	$1.479 \pm 0.022\,(\mathrm{stat}) \pm 0.037\,(\mathrm{sys})$
$[300-320[$	$1.203 \pm 0.017\,(\mathrm{stat}) \pm 0.015\,(\mathrm{sys})$
$[320-340[$	$1.244 \pm 0.016\,(\mathrm{stat}) \pm 0.026\,(\mathrm{sys})$
$[340-360[$	$1.409 \pm 0.019\,(\mathrm{stat}) \pm 0.015\,(\mathrm{sys})$
$[360-380[$	$1.448 \pm 0.022\,(\mathrm{stat}) \pm 0.020\,(\mathrm{sys})$
$[380-400[$	$1.472 \pm 0.033\,(\mathrm{stat}) \pm 0.014\,(\mathrm{sys})$
$[400-500[$	$1.487 \pm 0.017\,(\mathrm{stat}) \pm 0.012\,(\mathrm{sys})$
$[500-...]$	$1.505 \pm 0.014\,(\mathrm{stat}) \pm 0.004\,(\mathrm{sys})$

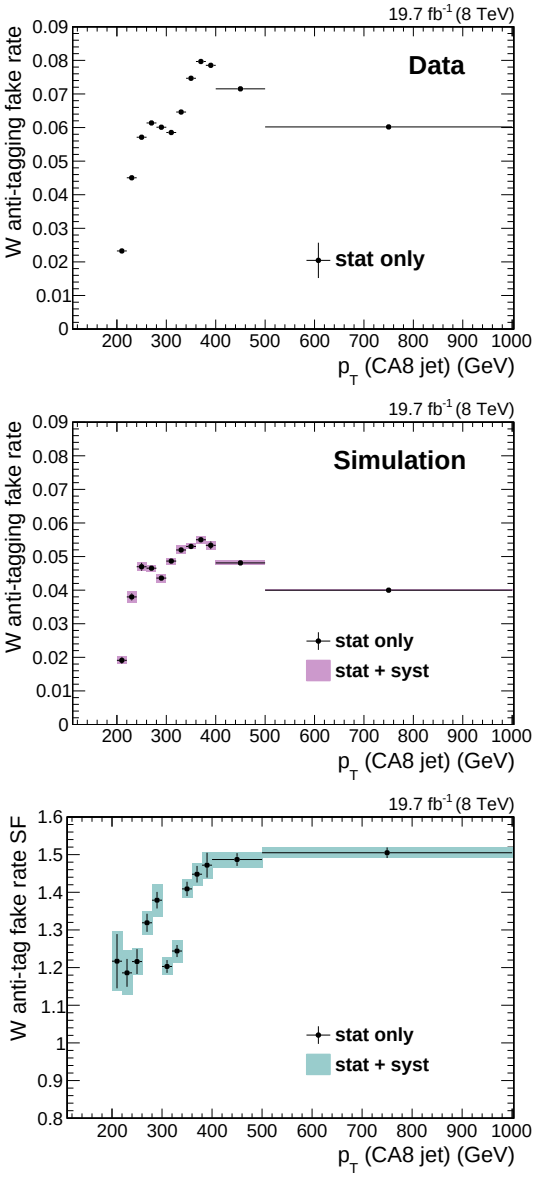

Figure 7.21: [top] Misidentification probability according to W anti-tagging for a CA8 jet versus jet p_{T} obtained from a multijet-enriched control region in data as described in the text. The shown uncertainties are statistical only. [middle] W boson anti-tag fake rate according to simulation. The uncertainty band includes statistical and systematic uncertainties. [bottom] Scale factor for W anti-tag fake rate versus CA8 jet p_{T} obtained from a multijet-enriched control region as described in the text. The uncertainty band includes statistical and systematic uncertainties.

Table 7.5: Summary of scale factors and their total uncertainty.

CA8 jet p_{T} (GeV)	$SF^{\mathrm{fake}}_{W\mathrm{tag}}$	$SF^{\mathrm{fake}}_{W\mathrm{masstag}}$	$SF^{\mathrm{fake}}_{W\mathrm{antitag}}$
$[200-220[$	1.04 ± 0.07	1.14 ± 0.06	1.22 ± 0.08
$[220-240[$	1.01 ± 0.04	1.12 ± 0.04	1.19 ± 0.06
$[240-260[$	1.16 ± 0.04	1.19 ± 0.03	1.22 ± 0.04
$[260-280[$	1.16 ± 0.03	1.25 ± 0.03	1.32 ± 0.04
$[280-300[$	1.15 ± 0.03	1.27 ± 0.03	1.38 ± 0.05
$[300-320[$	1.03 ± 0.02	1.13 ± 0.02	1.20 ± 0.03
$[320-340[$	1.14 ± 0.02	1.20 ± 0.03	1.24 ± 0.03
$[340-360[$	1.17 ± 0.02	1.30 ± 0.02	1.41 ± 0.03
$[360-380[$	1.18 ± 0.03	1.33 ± 0.02	1.45 ± 0.03
$[380-400[$	1.18 ± 0.03	1.34 ± 0.03	1.47 ± 0.04
$[400-500[$	1.15 ± 0.02	1.34 ± 0.02	1.49 ± 0.03
$[500-...]$	1.18 ± 0.02	1.37 ± 0.02	1.51 ± 0.02

7.5 Trigger and datasets

7.5.1 Data and trigger

The analysis presented in this thesis is based on $19.7\,\mathrm{fb}^{-1}$ of 8 TeV proton-proton collision data collected by the CMS experiment in 2012. The data is divided into four data taking periods (A, B, C, and D) to deal with changing conditions such as the average pileup or peak luminosity. Events used in the razor boost analysis are selected using two triggers from the CMS high level trigger system (see Section 5.2.7), requiring either the highest jet p_{T} or the scalar sum of jet transverse momenta, H_{T}, to be above certain thresholds. The jet p_{T} trigger, of which the different implementations are denoted by `PFJet*`, had a threshold on the p_{T} of the highest p_{T} jet of 320 GeV for most of the runs, and a threshold of 400 GeV for a short period. The H_{T}-based trigger, denoted by `PFHT*`, had an H_{T} threshold of 650 GeV. The two trigger algorithms were based on a fast implementation of the particle flow reconstruction method, which was described in Section 6.4.1. The exact names of the used trigger by run are given in Table 7.6. The primary datasets which include the data collected by these triggers are listed in Table 7.7.

In order to select events with unbiased jet p_{T} and H_{T} distributions, the trigger efficiency was measured from a sample of events collected with an orthogonal set of triggers requiring at least one electron or muon. The corresponding primary datasets are listed in Table 7.8. The trigger efficiency ϵ_{trig} is determined as a function of H_{T} and first jet p_{T}, and takes as basis for the measurement the baseline selection described further in Section 7.6.2,

$$\epsilon_{\mathrm{trig}} = \frac{\text{Events passing baseline and trigger selection}}{\text{Events passing baseline selection}}. \tag{7.75}$$

It was checked that the trigger efficiency obtained from either the electron or muon sample gives consistent results. Both samples were combined to derive

Table 7.6: Summary of HLT triggers that are used in this analysis. Events in a given run range are selected if they pass at least one of the listed triggers.

Period	Run range	PFJet HLT	PFHT HLT
Run2012A	190456 - 190738	PFJet320_v3	PFHT650_v5
	190762 - 191426	PFJet320_v4	PFHT650_v6
	191512 - 193686	PFJet320_v5	PFHT650_v7
Run2012B	193746 - 196027	PFJet320_v5	PFHT650_v8
	196039 - 197722	PFJet320_v5	PFHT650_v9
Run2012C	197770 - 199631	PFJet400_v6	PFNoPUHT650_v1
	199648 - 202585	PFJet320_v8	PFNoPUHT650_v3
	202807 - 203734	PFJet320_v9	PFNoPUHT650_v4
Run2012D	203754 - 208940	PFJet320_v9	PFNoPUHT650_v4

Table 7.7: List of primary datasets and corresponding run ranges, containing data for a total integrated luminosity of $19.712\,\mathrm{fb}^{-1}$.

Primary dataset	Run range
/Jet/Run2012A-22Jan2013-v1/	190456 - 193621
/HT/Run2012A-22Jan2013-v1/	190456 - 193621
/JetHT/Run2012B-22Jan2013-v1/	193833 - 196531
/JetHT/Run2012C-22Jan2013-v1/	198022 - 203742
/JetHT/Run2012D-22Jan2013-v1/	203777 - 208686

Table 7.8: List of primary datasets used to measure the trigger efficiency.

Primary dataset	Run range
/SingleElectron/Run2012A-22Jan2013-v1/	190456 - 193621
/SingleElectron/Run2012B-22Jan2013-v1/	193833 - 196531
/SingleElectron/Run2012C-22Jan2013-v1/	198022 - 203742
/SingleElectron/Run2012D-22Jan2013-v1/	203777 - 208686
/SingleMu/Run2012A-22Jan2013-v1/	190456 - 193621
/SingleMu/Run2012B-22Jan2013-v1/	193833 - 196531
/SingleMu/Run2012C-22Jan2013-v1/	198022 - 203742
/SingleMu/Run2012D-22Jan2013-v1/	203777 - 208686

the final trigger efficiency measurement in order to increase the statistical precision. Figure 7.22 shows, on the top plot, the trigger efficiency measured from data on the (H_{T}, first jet p_{T}) plane. We observe that the trigger is fully efficient for events with $H_{\mathrm{T}} > 800\,\mathrm{GeV}$. In order to account for the lower efficiency of the region with $H_{\mathrm{T}} < 800\,\mathrm{GeV}$, the measured trigger efficiency over the (H_{T}, first jet p_{T}) plane is applied as an event-by-event weight to the simulated samples. The bottom plot of Fig. 7.22 shows the effect of this trigger efficiency across the ($\mathrm{M_R}$,R^2) plane for the total simulated background. The uncertainty on the trigger efficiency is computed as the maximum of the statistical uncertainty, and the difference in trigger efficiency obtained using the baseline selection and no selection at all. As the uncertainties are not symmetric, we show both the up and down uncertainties in Fig. 7.23. They are generally below 5%.

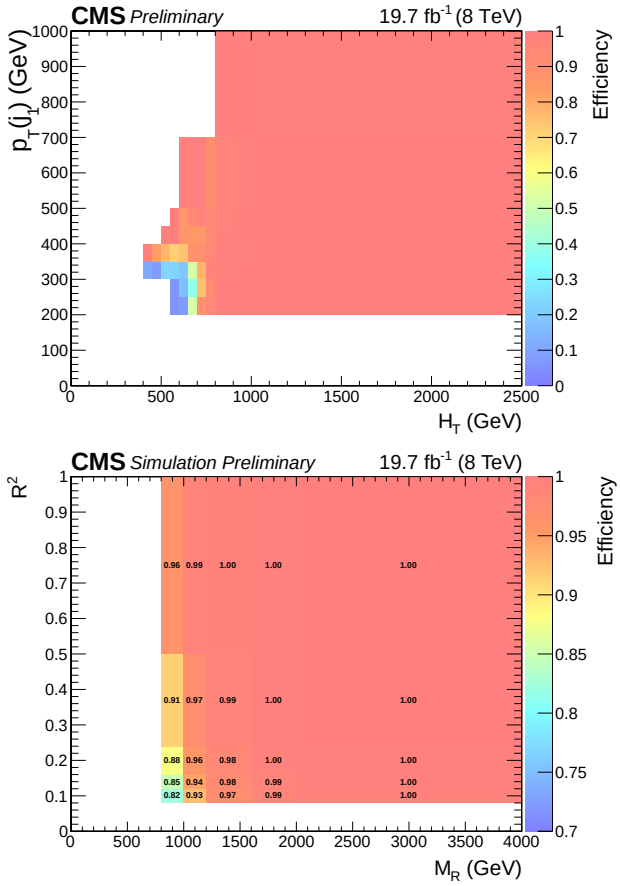

Figure 7.22: [top] The trigger efficiency, obtained from data, as a function of H_T and first jet p_{T} after the preselection mentioned in Section 7.6.2. [bottom] The trigger efficiency as a function of $\mathrm{M_R}$ and R^2 after the same preselection, obtained by applying the trigger efficiency as a function of H_{T} and first jet p_{T} to the total simulated background.

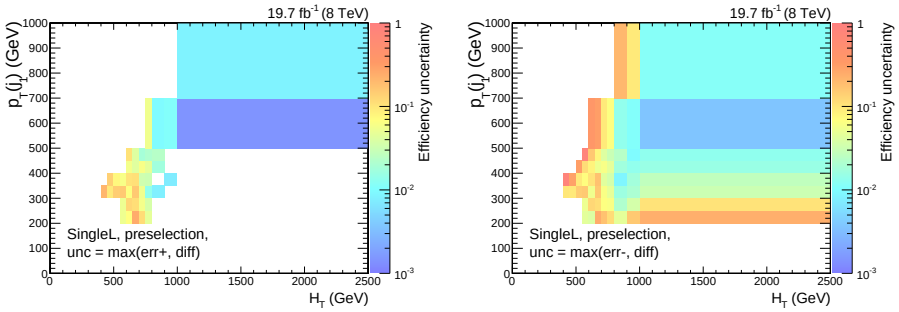

Figure 7.23: The magnitude of the up (left) and down (right) uncertainties on the efficiency as functions of H_{T} and first jet p_{T}, obtained using the combination of the SingleEle and SingleMu datasets.

7.5.2 Simulated samples

The simulated samples are used to investigate the characteristics of the background and signal processes. The complete list of simulated samples that are used in the analysis is given in Appendix A, in Tables A.1 and A.2 for the SM background samples, and in Table A.3 for the signal samples.

Multijet, $t\bar{t}$, $W(\to \ell\nu)$+jets, $\mathrm{Z}/\gamma^*(\to \ell\bar{\ell})$+jets, and $\mathrm{Z}(\to \nu\bar{\nu})$+jets events are generated using MadGraph 5.1.3.30 [89], as are the smaller backgrounds $W(\to q\bar{q})bb$, $WW\mathrm{Z}$+jets, $WW\gamma$+jets, ZZZ+jets, $t\bar{t}\gamma$+jets, and $t\bar{t}WW$+jets. Events for the $\mathrm{Z}/\gamma^*(\to c\bar{c})$, $\mathrm{Z}/\gamma^*(\to \mathrm{b\bar{b}})$, WW, $W\mathrm{Z}$, and ZZ processes are generated using pythia6.424 [90]. For all these samples CTEQ6L1 [173] is used as the set of parton distribution functions. Single top quark events are generated using powheg 1.0 [174, 175] with CT10 PDFs [176], and WWW, $W\mathrm{ZZ}$, $t\bar{t}W$ and $t\bar{t}\mathrm{Z}$ are generated using aMC@NLO [177] with CTEQ6M PDFs [173]. Signal events are produced using MadGraph 5.1.5.4 with CTEQ6L1 PDFs.

The parton level events are showered and hadronized using pythia6.426 with tune Z2* [178], except for the samples generated with aMC@NLO which use herwig [179, 180] for the parton shower and hadronization. The MadGraph samples are matched to the parton shower using the MLM technique, as discussed in Section 6.2.3. For the background events, the response of the CMS detector is simulated with a full simulation based on Geant4 [104]. A parameterized fast detector simulation, i.e. FastSim (Section 6.3.2), is used to simulate the detector response to the signal events.

7.6 Event selection

In this section I will discuss the event selection, starting by a review of how the razor variables will be used in the razor boost analysis. I will then detail the baseline selection that is applied to all events, before moving on to the signal region selection and the definition of the different control regions. Before we go ahead, it is useful to remind ourselves of the goal of the analysis and the basics

of the background estimation method. The aim of the analysis is to conduct a search for deviations from the SM in the high M_R-high R^2 region using hadronic events with at least one boosted W boson and one b jet. Standard Model backgrounds in the signal region are estimated via a likelihood method, using observations in three control regions in data, labelled Q, T and W, and transfer factors κ, calculated from simulated data, between these control regions and the signal region, S. The viability of the defined control regions to model the data in our signal region is tested by two closure tests, which are presented at the end of this section. For all details on the full background estimation itself, I refer to Section 7.7.

7.6.1 The razor variables in the razor boost analysis

As the name already indicates, the razor boost analysis uses the razor variables as the main discriminating variables to distinguish signal from background. They are well suited for the purpose of the analysis; they are designed to describe a signal due to pair production of heavy particles, each of which decays to a massless visible particle and a massive invisible particle, as is the case for the targeted models in the razor boost analysis. The signal will appear as a peak on a steeply falling background.

For clarity, the M_R and R^2 definitions from Section 7.3.4 are repeated here,

$$M_R = \sqrt{(E_{j_1} + E_{j_2})^2 - \left(p_z^{j_1} + p_z^{j_2}\right)^2}, \tag{7.76}$$

$$R^2 = \frac{M_T^R}{M_R}, \tag{7.77}$$

$$\text{with } M_T^R = \sqrt{\frac{|\vec{E}_T^{\text{miss}}|}{2}\left(p_T^{j_1} + p_T^{j_2}\right) - \frac{\vec{E}_T^{\text{miss}}}{2} \cdot \left(\vec{p}_T^{\,j_1} + \vec{p}_T^{\,j_2}\right)}, \tag{7.78}$$

where we have replaced the relevant variables with their megajet counterparts (denoted with indices j_i), and we have defined p_T as the magnitude of the transverse momentum vector $\vec{p}_T$.

The two-dimensional (2D) distributions of R^2 versus M_R for both background and an example signal model are shown in Fig. 7.24. A very loose selection is applied, requiring the presence of a good primary vertex and at least three jets, one of which should have $p_T > 200\,\text{GeV}$. It is clear that the background, dominated by QCD multijet production, is located in the low M_R and low R^2 regions, and falls off steeply when M_R or R^2 is increased. The signal, on the other hand, shows a peaking behaviour in M_R and is located substantially higher in the (M_R,R^2) space. The M_R distribution peaks at higher values for larger sparticle masses and higher mass differences between the gluino and the LSP. This is illustrated in Figs. 7.25 and 7.26, which show the M_R and R^2 distributions for several signal points using the same loose selection as defined above. The width of the M_R distribution also increases with the gluino mass. This is why we use an increasingly larger bin width for M_R when defining our analysis regions.

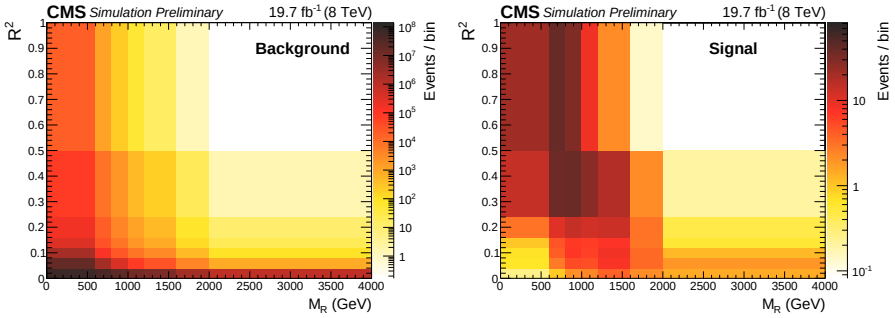

Figure 7.24: Distribution of the overall SM backgrounds and a T1ttcc signal with $m_{\tilde{g}} = 1\,\text{TeV}$, $m_{\tilde{t}} = 325\,\text{GeV}$ and $m_{\tilde{\chi}_1^0} = 300\,\text{GeV}$, both obtained from MC, on the (M_R,R^2) space. A very loose selection is used: a good primary vertex and at least three jets, one of which should have $p_T > 200\,\text{GeV}$.

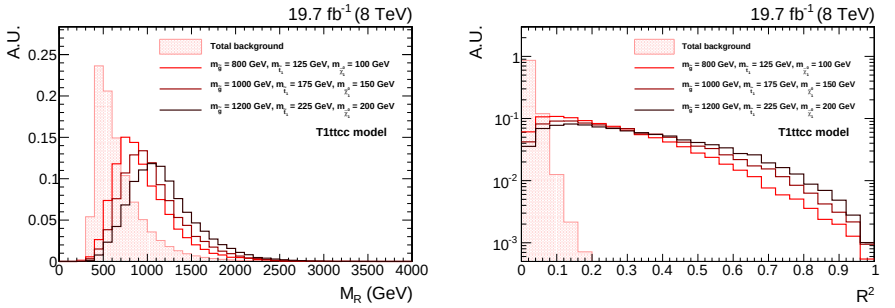

Figure 7.25: Distribution of M_R (left) and R^2 (right) for the total SM background, and several T1ttcc signal points. A very loose selection is used, including only the requirement of having a good primary vertex and at least three jets, one of which with $p_T > 200\,\text{GeV}$.

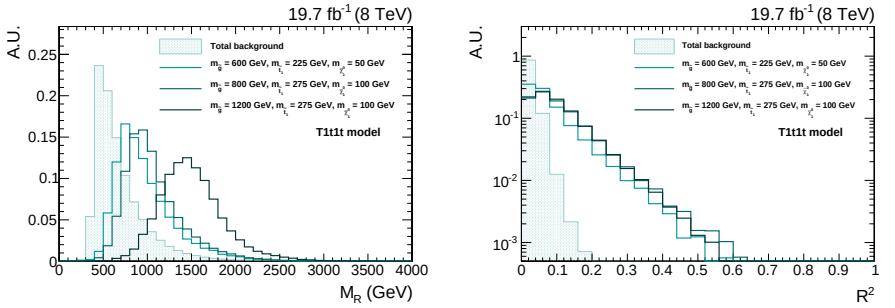

Figure 7.26: Distribution of M_R (left) and R^2 (right) for the total SM background, and several T1t1t signal points. A very loose selection is used, including only the requirement of having a good primary vertex and at least three jets, one of which with $p_T > 200\,\text{GeV}$.

In order to retain sensitivity to low $E_{\mathrm{T}}^{\mathrm{miss}}$ scenarios, we choose to reduce the minimal R^2 requirement compared to previous razor analyses [157,158,181], while simultaneously raising the $\mathrm{M_R}$ requirement to access the boosted phase space, and keep the background to a manageable level. As explained before, $\mathrm{M_R}$ is a measure of mass difference, and since high boost requires high mass differences, we opt to work in the high $\mathrm{M_R}$ region to obtain better expected signal significance. The baseline requirement implemented in this analysis is that events fall within the ranges

- $\mathrm{M_R} > 800\,\mathrm{GeV}$,
- $\mathrm{R}^2 > 0.08$,

providing a good balance between background suppression and signal acceptance. This part of the razor plane is further divided into 25 bins, whose edges are clear from Fig. 7.24.

7.6.2 Baseline selection

The baseline selection for the razor boost analysis is driven by the two components of its name, in addition to requiring the event to be of good quality. As we use the razor variables, we need to be able to compute them. This means that there should be at least two jets in the final state. The megajets from which the razor variables are computed, are constructed from the AK5 jets (see Table 6.3 for their definition). After sensitivity studies, it was decided to raise the requirement on the jet multiplicity from two to three, reducing the background, while maintaining very good signal efficiency. The signal processes that are the main focus of this analysis usually have even more reconstructed jets, as can be seen from Fig. 7.27. In order to remain as inclusive as possible, we did not, however, raise this threshold further. The minimal requirements on the razor variables themselves are, as mentioned before, $\mathrm{M_R} > 800\,\mathrm{GeV}$ and $\mathrm{R}^2 > 0.08$. This selection is complementary to that of previous razor analyses. By requiring a larger minimal $\mathrm{M_R}$, consistent with the boosted scenario, we can explore the low R^2 region, which is important for signals with more compressed mass spectra. Access to the boosted phase space is also provided by making the requirement that at least one AK5 jet satisfies $p_{\mathrm{T}} > 200\,\mathrm{GeV}$. In summary, events are required to satisfy the following baseline selection:

1. Satisfy all detector- and beam-related cleaning filters (Section 6.4.3) to remove events with high energy and high transverse momentum imbalance,
2. Have at least one primary vertex satisfying quality cuts on vertex fit and position with respect to the detector centre,
3. Have at least three selected AK5 jets of which at least one has $p_{\mathrm{T}} > 200\,\mathrm{GeV}$, thereby defining the boosted phase space,
4. Satisfy $\mathrm{M_R} > 800\,\mathrm{GeV}$ and $\mathrm{R}^2 > 0.08$ (where the megajets are constructed from the selected AK5 jets).

In addition to these requirements, we also impose the trigger conditions. For data events, we require that one of the triggers listed in Table 7.6 was fired. For simulated events, we apply an event-by-event trigger efficiency, as explained in Section 7.5.1.

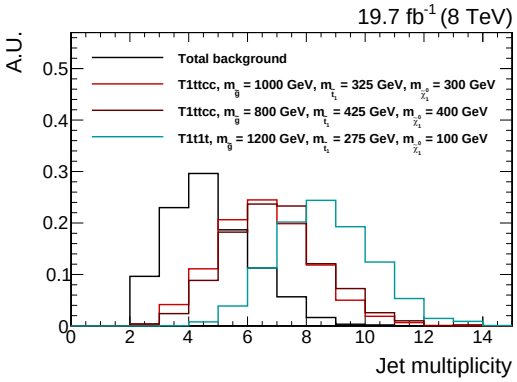

Figure 7.27: Distribution of the number of jets for the total SM background and three example signal points in the signal region. The full selection except for the jet multiplicity requirement was applied. The histograms are normalized to unit area.

In Fig. 7.28 we compare the observed data to the simulation for events passing this baseline selection. As will be the case for all following comparisons of data versus simulation, the simulation is scaled to the expected number of events according to the (NN)LO cross section of each process. In addition, the simulated events are reweighted in order to correct for mismodelling of the pileup, trigger, b tagging, top p_{T} distribution, etcetera. Most of these reweightings are performed for each region, but for some, such as the b tagging scale factors, the reweighting is only applied when the source of the reweighting is used explicitly in the selection. For the b tagging this would be a specific requirement or veto on the b-tagged jet multiplicity. Hence, the reweighting is not applied in the baseline selection region. A full overview of these sources of event reweighting is given in Table 7.9. Each of these sets of event weights has an associated uncertainty. These will be taken into account as systematic uncertainties on the final background prediction, as explained in more detail in Section 7.8. From the figure, we see that there is good agreement between data and simulation for the highest $\mathrm{M_R}$ and R^2 bins, i.e. when the contribution of QCD multijet MC is very small. For the first $\mathrm{M_R}$ and R^2 bins, which are dominated by multijet production, the simulation underpredicts the data by about 50%.

The data/MC comparison is also displayed in table form. The first section of Tables 7.12 and 7.13 shows the expected number of events for the different background processes and observed number of events in data for several steps in the baseline cutflow. The entry listed as "No selection" corresponds to the total number of events expected when no selection is applied. It is equal

Table 7.9: Sources of event reweighting, when they are applied, and to which process.

Source	Region	Process
Pileup	All	All
Trigger	All	All
b tagging	If b tagging used	All
W tagging	If W tagging used	All
W mass-tagging	If W mass-tagging used	All
W anti-tagging	If W anti-tagging used	All
Top p_{T} spectrum	All	$t\bar{t}$
Initial state radiation	All	Signal

to the cross section of the process times the integrated luminosity. The row corresponding to "$n_{PV} > 0$" gives the event counts after applying the cleaning filters, the relevant event reweightings, and the requirement that there be at least one good primary vertex. The background composition after the full baseline selection, expressed in percentages, is reported in Table 7.11.

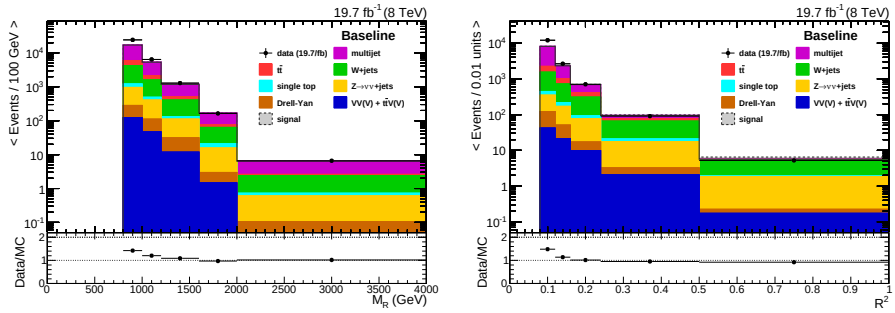

Figure 7.28: Comparison between data and simulation for the $\mathrm{M_R}$ (left) and R^2 (right) distribution in the baseline selection region. The bin entries are scaled proportional to the bin width.

7.6.3 Signal region selection

The signal region selection aims for a good discrimination between possible signal processes and the SM backgrounds. As mentioned before, the signals we target with this search have b tagged jets and boosted W bosons in the final state. Therefore, we require, on top of the baseline selection, the presence of at least one CSV medium b tagged jet, and at least one W boson tagged jet. AK5 jets are used for b tagging, whereas for W tagging we use the CA8 jets, as explained in Section 7.4. Additionally, we only consider fully-hadronic events and thus select only those events without loose electrons or muons,

and no isolated tracks. These selection criteria already reduce the background substantially, but the achieved signal separation is not yet sufficient. We need an additional handle on the QCD multijet production, which is the dominant background at this stage.

Missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, in multijet events is largely due to jet mismeasurements, rather than the escape of weakly interacting particles, such as neutrinos or the neutralinos in signal events. The $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector will, therefore, often be aligned with one of the jets. Based on this we can expect that $\Delta\phi_{min}$, the minimum of the angles between $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse momentum of the leading three jets, will be a good discriminant between multijet events and events with real $E_{\mathrm{T}}^{\mathrm{miss}}$.

$$\Delta\phi_{min} = \min_{i=1,2,3} \Delta\phi(\vec{E}_{\mathrm{T}}^{\mathrm{miss}}, \vec{p}_{\mathrm{T}}^{\,i}), \tag{7.79}$$

where i runs over the three leading AK5 jets. We require $\Delta\phi_{min} > 0.5$ to suppress multijet events. The $\Delta\phi_{min}$ distribution, obtained from simulation, before applying this selection is shown in Fig. 7.29. The multijet events are clearly gathered in the first few bins, while the signal extends to much higher values.

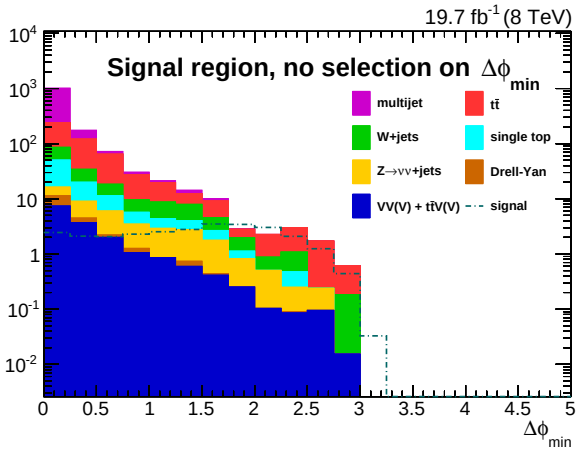

Figure 7.29: Simulated $\Delta\phi_{min}$ distribution with all signal region requirements applied except $\Delta\phi_{min} > 0.5$. QCD multijet events are clearly gathered in the first few bins. An example signal point is overlaid, and is seen to have a much flatter distribution, extending to high $\Delta\phi_{min}$ values.

A summary of the signal selection is presented in Table 7.10. Figure 7.30 shows the simulated distributions in the signal region for the $\mathrm{M_R}$ and R^2 variables. The number of events in simulation and data, and the background composition in percent, are reported in Tables 7.12 and 7.13, and Table 7.11, respectively. The signal region is $t\bar{t}$ dominated, with additional contributions from $W(\to \ell\nu)$+jets and multijet processes. The p_{T} distribution of the highest p_{T} tagged W boson jet in the event is shown in Fig. 7.31, alongside the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution.

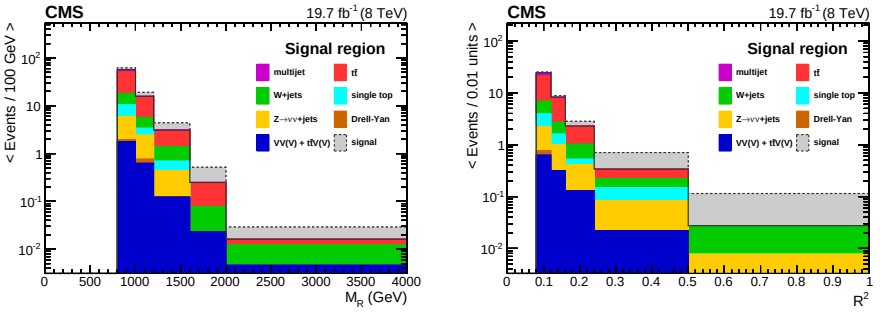

Figure 7.30: Simulated M_R (left) and R^2 (right) distributions in the signal region. An example signal point, corresponding to the T1ttcc mass point with $m_{\tilde{g}} = 1\,\text{TeV}$, $m_{\tilde{t}_1} = 325\,\text{GeV}$ and $m_{\tilde{\chi}_1^0} = 300\,\text{GeV}$, is stacked on top of the background processes. The bin entries are normalized proportional to the bin width.

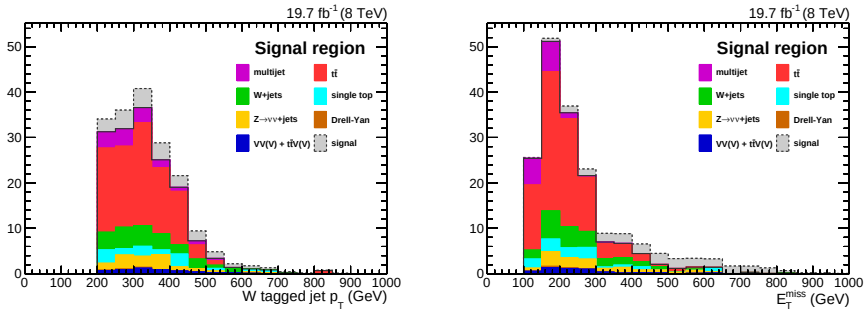

Figure 7.31: Simulated W tagged jet p_{T} (left) and $E_{\text{T}}^{\text{miss}}$ (right) distributions in the signal region. An example signal point, corresponding to the T1ttcc mass point with $m_{\tilde{g}} = 1\,\text{TeV}$, $m_{\tilde{t}_1} = 325\,\text{GeV}$ and $m_{\tilde{\chi}_1^0} = 300\,\text{GeV}$, is stacked on top of the background processes.

7.6.4 Control region selection

The control regions are defined such that they are enriched in one of the three main backgrounds in the signal region: $t\bar{t}$, QCD multijet, and $W(\to \ell\nu)$+jets. The control region selections will be as close as possible to the signal selection, while maximizing the purity in a given background process and minimizing possible signal contamination. To construct the control regions, different requirements on the multiplicity of leptons, b tagged jets, and W tagged jets are placed. In the following subsections each control region will be discussed in detail.

T region

The dominant background in the signal region is $t\bar{t}$ production. We define the T region as a dedicated control region enriched in this process as well as single top production. Similarly to the signal, $t\bar{t}$ and single top events are expected to have b jets in the final state, as well as real hadronically decaying W bosons. We will thus not change the b tagging and W tagging aspects of the signal selection when defining the T region.

The main change we do make, is requiring the presence of exactly one loose electron or muon, while removing the requirement that there be no isolated track. This selection is already sufficient to obtain a region with a good purity of $t\bar{t}$ events. There is, however, a possibility of substantial signal contamination. To address this issue, we place an upper boundary on the value of the transverse mass, m_{T}, computed from the lepton transverse momentum and $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$,

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}}(1-\cos\Delta\phi)}, \tag{7.80}$$

with $\Delta\phi$ the difference in azimuthal angle between lepton and $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$. The distribution of m_{T} exhibits a kinematic edge at the mass of the W boson for the $t\bar{t}$ process, an edge not present for signal events due to the extra contribution to the $E_{\mathrm{T}}^{\mathrm{miss}}$ from the invisible neutralinos. We require $m_{\mathrm{T}} < 100\,\mathrm{GeV}$ in order to reduce signal contamination and retain most of the $t\bar{t}$ contribution. The m_{T} distribution at this selection level is shown in Fig. 7.32. It can be seen that signal events could populate the tails of this variable.

The final ingredient for the full T region selection, is the $\Delta\phi_{min} > 0.5$ requirement. This is applied in order to obtain a selection, and thus kinematic properties, as close as possible to the signal region. The $\Delta\phi_{min}$ distribution in the T region is shown in Fig. 7.33. A summary of the full T region selection is presented in Table 7.10.

Tables 7.12 and 7.13, in particular the line listing T in the first column, show a breakdown of the contributions of the various background components, as determined directly from simulation, as well as the observed data counts, and expected counts for an example signal point. From Table 7.11 we see that the obtained purity in the T region is about 83% for $t\bar{t}$ and single top processes combined. This is also illustrated in the $\mathrm{M_R}$ and R^2 distributions, as shown on Fig. 7.34. There is good agreement between data and simulation in the T region. Any discrepancies can be accommodated by the systematic

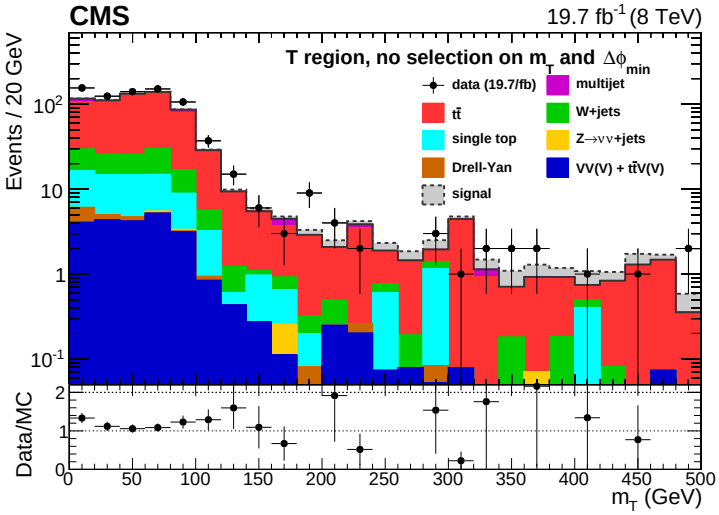

Figure 7.32: Comparison between data and simulation of the m_{T} distribution in the T region without any selection on m_{T} or $\Delta\phi_{min}$. An example signal point, corresponding to the T1ttcc mass point with $m_{\tilde{g}} = 1\,\mathrm{TeV}$, $m_{\tilde{t}_1} = 325\,\mathrm{GeV}$ and $m_{\tilde{\chi}^0_1} = 300\,\mathrm{GeV}$, is stacked on top of the background processes. The kinematic edge at the W boson mass is clearly visible for the SM backgrounds, whereas the example signal point extends out to high m_{T}.

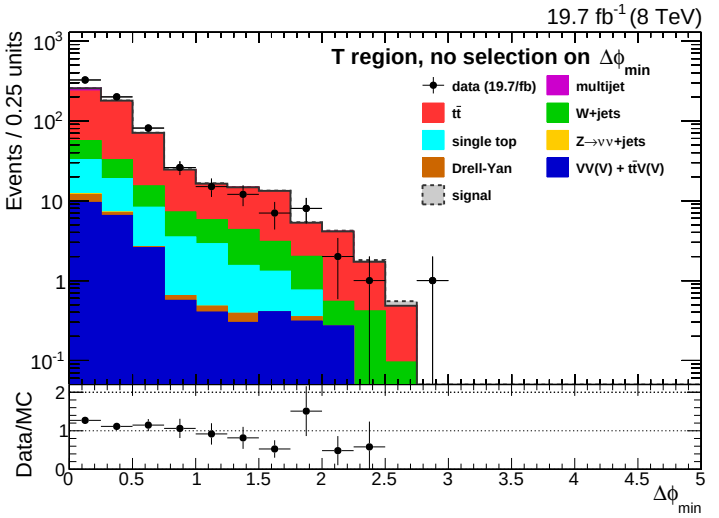

Figure 7.33: Comparison between data and simulation of the $\Delta\phi_{min}$ distribution in the T region without any selection on $\Delta\phi_{min}$. An example signal point, corresponding to the T1ttcc mass point with $m_{\tilde{g}} = 1\,\mathrm{TeV}$, $m_{\tilde{t}_1} = 325\,\mathrm{GeV}$ and $m_{\tilde{\chi}^0_1} = 300\,\mathrm{GeV}$, is stacked on top of the background processes. Because the $m_{\mathrm{T}} < 100\,\mathrm{GeV}$ requirement has already been applied, there is only a very minimal signal contamination that remains in this region.

uncertainties on the MC simulation. Given that we will only use the simulation in ratios between different regions, any small discrepancies will not affect the analysis. We also compare the shape of the M_R and R^2 distributions in the T region versus the S region. We observe from Fig. 7.35 that the shapes are very similar. This is important as we will use global transfer factors κ to translate between the T and S regions.

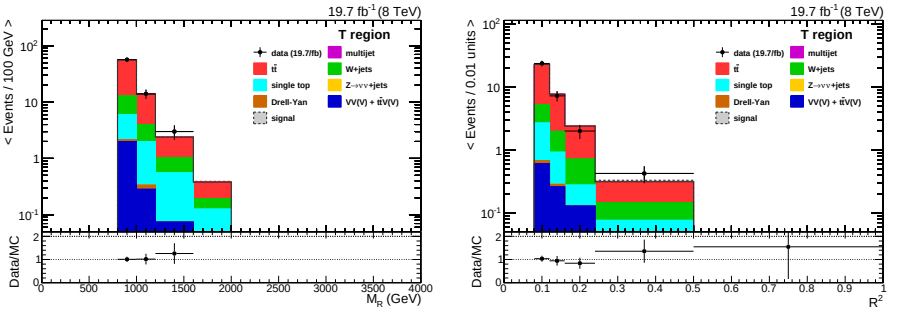

Figure 7.34: Comparison between data and simulation of the M_R (left) and R^2 (right) distribution in the T control region.

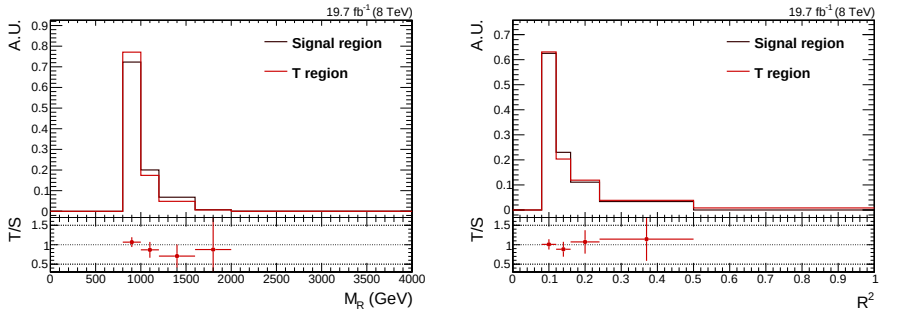

Figure 7.35: Comparison of the shape of the M_R (left) and R^2 (right) distribution for the $t\bar{t}$ simulation in the T region versus the S region. The histograms are normalized to unit area.

W region

The second largest background in the signal region comes from $W(\rightarrow l\nu)$+jets production, where the lepton from the W boson decay is lost. Because the W boson decays leptonically, any events passing the signal selection necessarily contain a q/g initiated CA8 jet that is misidentified as a boosted W boson. We do not expect hadronically decaying W bosons to contribute to the background in the signal region, because those events would have very small missing energy, in addition to not having real b jets. Even though we do not explicitly require a minimal $E_{\mathrm{T}}^{\mathrm{miss}}$, the R^2 variable is highly correlated with $E_{\mathrm{T}}^{\mathrm{miss}}$, and the applied

selection on R^2 effectively induces a minimal requirement of around 100 GeV on E_T^{miss}, as can be seen in Fig. 7.31. This expectation was verified by checking the contribution of $W(\rightarrow q\bar{q}')b\bar{b}$ using simulation, where it was found that no events pass our signal selection.

In order to accurately model the $W(\rightarrow l\nu)$+jets background, we define a dedicated control region, the W region, as similar to the signal region as possible. As for the T region, we require the presence of a loose electron or muon. Additionally, we also veto any event that contains a CSV loose b tagged jet. Since for the $W(\rightarrow l\nu)$+jets process, there is no real hadronically decaying W boson, we opt not to use the full W tagger. In order to increase the number of events available in the control region, we require instead the presence of at least one W boson mass-tagged jet. This ensures that we are in a kinematically similar phase space, because we keep the jet mass requirement, but prevents the penalty from requiring a two-prong decay for a process for which these decays do not occur.

We also require $m_T < 100$ GeV in order to reduce possible signal contamination. Unlike for the T region, we additionally require $m_T > 30$ GeV in order to reduce remaining contamination from multijet events in the W region. Those events generally have low E_T^{miss}, which translates in low m_T. Multijet contamination in the T region is negligible because of the requirement that there be at least one jet tagged a coming from a b quark. Finally, we also require $\Delta\phi_{min} > 0.5$, as was the case for the signal region. A summary of the W region selection is presented in Table 7.10.

Figure 7.36 shows the m_T distribution in the W region before making any selection m_T or on $\Delta\phi_{min}$. The $\Delta\phi_{min}$ distribution in the W region without applying a selection on $\Delta\phi_{min}$ is given in Fig. 7.37. The full breakdown of the backgrounds in this region is listed in Tables 7.12 and 7.11, as was done for the T region. The W region is about 85% pure in the $W(\rightarrow l\nu)$+jets process.

A comparison between data and simulation for the M_R and R^2 distributions is shown in Fig. 7.38. A good agreement is observed. The offset in the normalization will be absorbed automatically during the background estimation procedure. We also compare the shapes of the M_R and R^2 distributions for the $W(\rightarrow \ell\nu)$+jets simulation in the W region versus the S region. From Fig. 7.39 we observe that there is good agreement, within statistical uncertainties.

Q region

The final background for which we define a control region, is QCD multijet production. To define a region Q enriched in multijet production, we start from the baseline selection, and add the requirement that there be no loose lepton or isolated track present, as is the case for the signal region selection. We also veto events containing a CSV loose b tagged jet in order to veto contamination of top quarks.

Multijet events do not produce W bosons; any CA8 jet that passes the W boson tagger, is thus by definition a misidentified W boson jet. To reach a similar kinematic phase space as the signal region, we require that there be at least one W boson anti-tagged jet. The W boson anti-tagging still requires

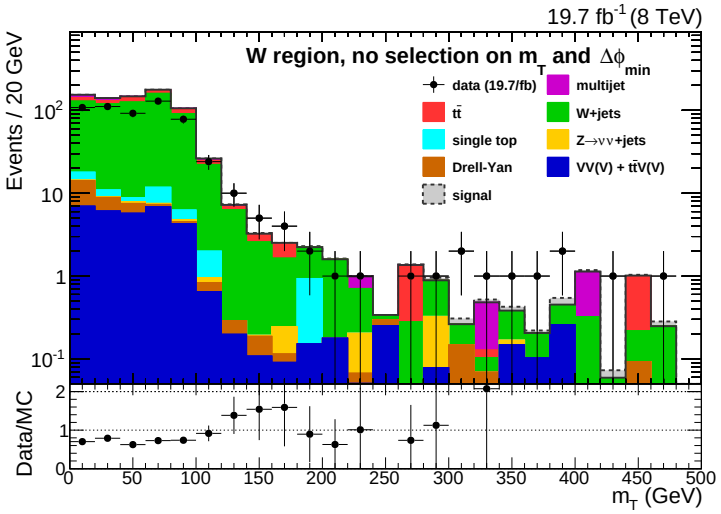

Figure 7.36: Comparison between data and simulation of the m_{T} distribution in the W region without any selection on m_{T} or on $\Delta\phi_{min}$. An example signal point, corresponding to the T1ttcc mass point with $m_{\tilde{g}} = 1\,\mathrm{TeV}$, $m_{\tilde{t}_1} = 325\,\mathrm{GeV}$ and $m_{\tilde{\chi}^0_1} = 300\,\mathrm{GeV}$, is stacked on top of the background processes.

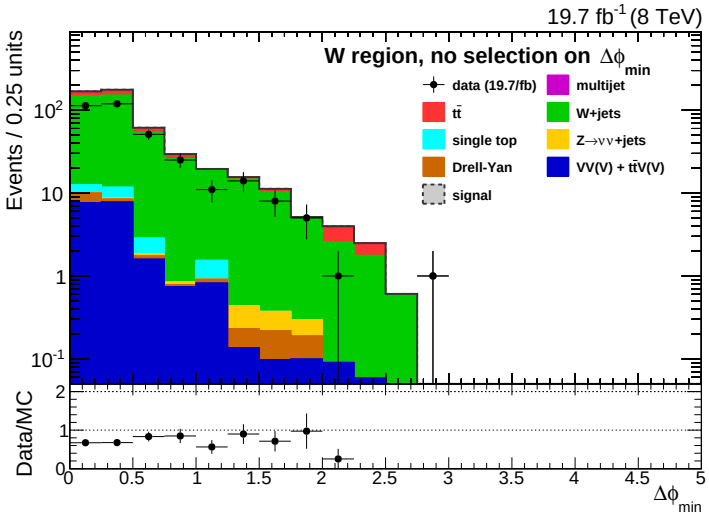

Figure 7.37: Comparison between data and simulation of the $\Delta\phi_{min}$ distribution in the W region without any selection on $\Delta\phi_{min}$.

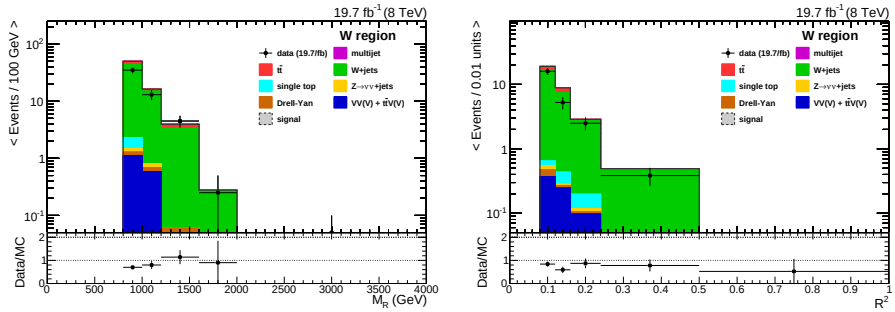

Figure 7.38: Comparison between data and simulation of the M_R (left) and R^2 (right) distributions in the W region.

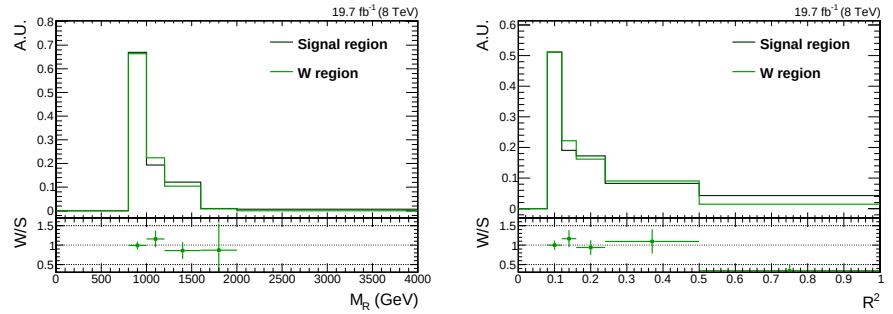

Figure 7.39: Shape of the M_R (left) and R^2 (right) distribution for $W(\to \ell\nu)$+jets simulation in the W region versus the S region.

that the jet mass is consistent with the W boson mass, but inverts the N-subjettiness requirement. As we do not expect QCD multijet events to have the two-prong jet-substructure of actual boosted W bosons, this will enhance the statistical power and purity of the Q region.

As explained before, QCD multijet events are expected to have small values for the $\Delta\phi_{min}$ variable. This can be seen from Fig. 7.40, where the multijet events are clearly located in the first couple bins. For the Q region definition we will thus reverse and tighten the $\Delta\phi_{min}$ requirement from the signal region. We require $\Delta\phi_{min} < 0.3$, which results in a very good purity of better than 90% according to simulation. The full breakdown of the backgrounds according to simulation can again be found in Tables 7.12 and 7.11, this time in the line corresponding to Q. Figure 7.41 shows a comparison between data and simulation for the M_R and R^2 distributions. It is not surprising to observe that the agreement between data and simulation is not very good, as this is a known feature of the QCD multijet simulation with normalization to LO cross sections. Since we only use ratios of simulated counts in the background estimation method, and the proportion of multijet events in the signal region is expected to be small, the observed mismodelling will, fortunately, not affect our analysis greatly.

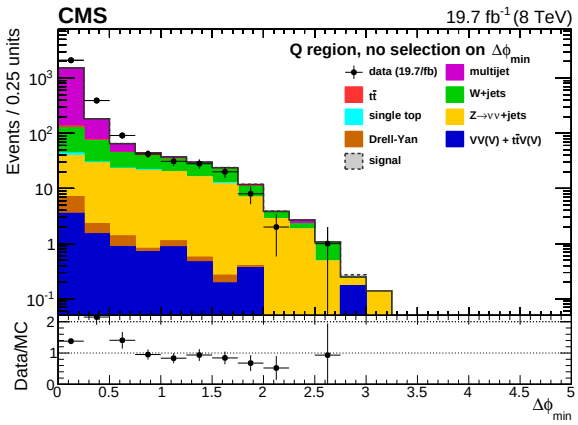

Figure 7.40: Comparison between data and simluation of the $\Delta\phi_{min}$ distribution in the Q region before any selection on $\Delta\phi_{min}$.

In Fig. 7.42 we show a comparison between the shapes of the M_R and R^2 distributions for the QCD multijet simulation in the signal region versus the Q region. Given the limited statistical precision for the multijet simulation in the signal region, it is hard to draw any conclusions from this comparison. Therefore, we also show in Fig. 7.43 the comparison for the Q and S regions without selecting a particular $\Delta\phi_{min}$ region. The shape difference in this region is quite small. As will be explained further in Section 7.6.5, we will assign a 40% systematic uncertainty on the multijet transfer factors in the background prediction to account for the possible shape difference induced by the $\Delta\phi_{min}$ cut.

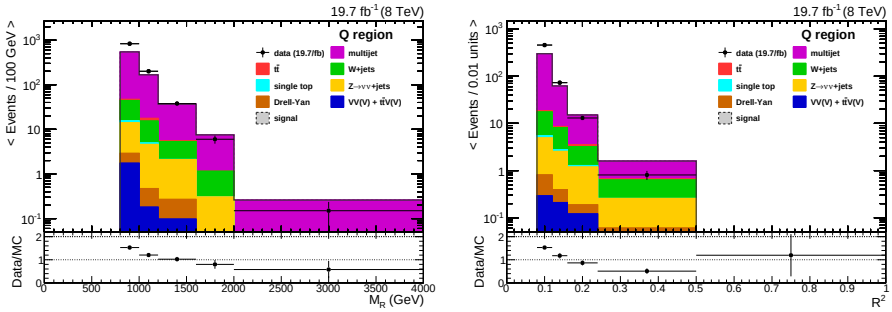

Figure 7.41: Comparison between data and simulation of the M_R (left) and R^2 (right) distributions in the Q region.

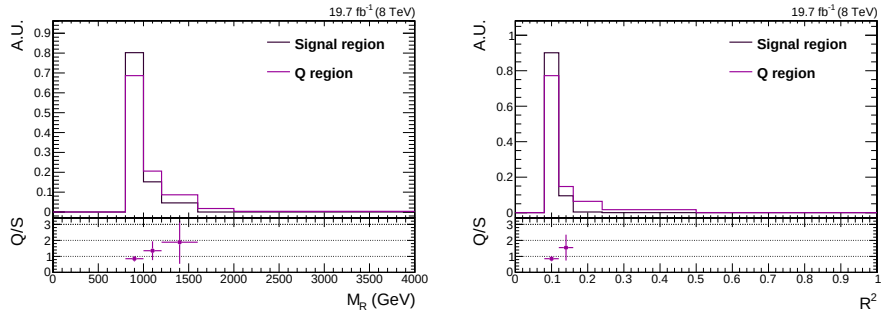

Figure 7.42: Shape of the M_R (left) and R^2 (right) distribution for the QCD multijet simulation in the Q region versus the S region.

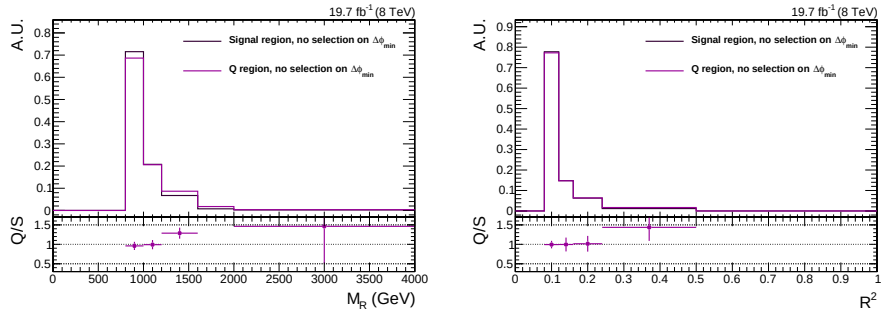

Figure 7.43: Shape of the M_R (left) and R^2 (right) distribution for the QCD multijet simulation in the Q region versus the S region with the requirement on $\Delta\phi_{min}$ removed in both regions.

Table 7.10: Summary of selection for signal (S) and control regions (Q, T, W). The selections used to define validation regions (S', Q') used in the closure tests are also included for completeness. The requirements listed in this table are to be applied in addition to those corresponding to the baseline selection.

Selection	S	S'	Q	Q'	T	W
nr of b jets	≥ 1	≥ 1	0	0	≥ 1	0
nr of mass-tagged W's	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1
nr of tagged W's	≥ 1	≥ 1	-	-	≥ 1	-
nr of anti-tagged W's	-	-	≥ 1	≥ 1	-	-
nr of loose leptons	0	0	0	0	1	1
nr of isolated tracks	0	0	0	0	-	-
m_T	-	-	-	-	$< 100\,\text{GeV}$	30–100 GeV
$\Delta\phi_{min}$	> 0.5	< 0.5	< 0.3	> 0.5	> 0.5	> 0.5

Table 7.11: Background composition according to simulation

Selection	Multijet	$t\bar{t}$	$W(\to \ell\nu)$+jets	Single top	$Z(\to \nu\bar{\nu})$+jets	Diboson	Other
Baseline	62.8%	10.2%	18.7%	1.6%	4.6%	0.5%	1.6%
S	9.2%	56.3%	14.4%	7.3%	7.9%	2.3%	2.6%
Q	90.2%	0.7%	5.8%	0.2%	2.4%	0.2%	0.3%
T	0.0%	73.9%	13.3%	8.8%	0.0%	1.3%	2.7%
W	0.0%	10.3%	84.8%	1.1%	0.4%	2.4%	1.0%
Q'	12.3%	2.8%	36.8%	1.0%	45.0%	1.7%	0.4%
S'	69.5%	20.3%	2.8%	3.8%	0.8%	0.4%	2.4%

Table 7.12: Cutflow table, event counts are normalized to $19.7\,\mathrm{fb}^{-1}$. The row corresponding to "$n_{PV} > 0$" gives the event counts after applying the cleaning filters, pileup reweighting, top p_{T} reweighting for $t\bar{t}$, ISR reweighting for signal, and the requirement of at least one good primary vertex. The cross sections used for each sample are listed in the second row.

Selection	Multijet 10.4×10^7 pb	$t\bar{t}$ 245.8 pb	$W\to\ell\nu$+jets 111.5 pb	Diboson 95.4 pb	Single top 114.9 pb	Z$\to\nu\nu$+jets 588.3 pb	DY$\to\ell\ell$+jets 22.6 pb	Triboson 0.69 pb	$t\bar{t}V$ 1.88 pb
No selection	2.1×10^{11}	4.9×10^6	2.2×10^6	1.9×10^6	2.3×10^6	1.2×10^7	4.5×10^5	1.2×10^4	3×10^4
$n_{PV} > 0$	1.05×10^{11}	4.42×10^6	2.02×10^6	1.08×10^6	1.72×10^6	2.87×10^6	3.7×10^5	8.46×10^3	2.6×10^4
$n_j \geq 3$	2.04×10^{10}	4.08×10^6	1.51×10^6	5.19×10^5	1.10×10^6	6.24×10^5	3.06×10^5	5.64×10^3	2.49×10^4
$p_{\mathrm{T}}(j_1) > 200\,\mathrm{GeV}$	1.82×10^8	2.88×10^5	4.36×10^5	1.86×10^4	6.08×10^4	5.89×10^4	6.61×10^4	924	5.24×10^3
$M_R > 800$, $R^2 > 0.08$	3.47×10^4	5.83×10^3	1.17×10^4	309	900	3.25×10^3	422	40.2	183
Trigger	3.15×10^4	5.12×10^3	9.38×10^3	249	786	2.32×10^3	367	36.4	166
no lepton	3.09×10^4	1.87×10^3	3.75×10^3	96.3	311	2.30×10^3	145	12.6	58.5
$n_b \geq 1$	9.37×10^3	1.51×10^3	590	25.2	226	302	29.0	4.48	46.3
$n_W \geq 1$	841	332	56.4	8.52	56.7	22.1	5.28	1.98	9.68
S	14.8	90.4	23.1	3.7	11.7	12.7	0.59	0.98	2.6
S'	826	242	33.3	4.8	45	9.48	4.7	1	7.06
$n_b = 0$	1.25×10^4	98.3	1.70×10^3	35.6	25.9	1.25×10^3	46.5	4.19	3.56
$n_{aW} \geq 1$	1519	18.7	204	8.36	7.40	158	5.41	0.751	0.819
Q	1447	10.6	93.1	3.88	3.94	38.9	3.68	0.28	0.52
Q'	27	6.1	81	3.76	2.31	99.1	1.05	0.337	0.171
1 lepton	585.9	2.74×10^3	5.52×10^3	132	421	22.1	164	19.2	88.5
$n_b \geq 1$	236.7	2.17×10^3	625	29.9	301	4.14	28.7	5.36	68.3
$n_W \geq 1$	24.3	496	61.6	10.0	50.9	0.56	3.57	2.36	16.0
T	0	112	20.2	2.0	13.3	0	0.38	0.50	3.2
$n_b = 0$	150.5	153	2.86×10^3	52.8	41.3	11.5	55.8	7.05	5.94
$n_Y \geq 1$	30.8	79.1	605	33.1	13.8	2.4	13.1	4.57	2.61
W	0	15.5	127	3.6	1.6	0.64	0.59	0.52	0.29

Table 7.13: Cutflow summary table, event counts are normalized to $19.7\,\mathrm{fb}^{-1}$. The signal is the $m_{\tilde{g}} = 1000\,\mathrm{GeV}$, $m_{\tilde{t}_1} = 325\,\mathrm{GeV}$, $m_{\tilde{\chi}_1^0} = 300\,\mathrm{GeV}$ point of the T1ttcc scan. The row corresponding to "$n_{PV} > 0$" gives the event counts after applying the cleaning filters, pileup reweighting, top p_T reweighting for $t\bar{t}$, ISR reweighting for signal, and the requirement of at least one good primary vertex. The column indicating the total number of events also includes some smaller processes that only contribute at the early stages of the event selection, and which were omitted in Table 7.12. The cross section used for the signal sample is listed in the second row.

Selection	Total	Signal 0.02435 pb	Data
No selection	2.1×10^{11}	499	
$n_{PV} > 0$	1.05×10^{11}	479	
$n_j \geq 3$	2.05×10^{10}	472	
$p_\mathrm{T}(j_1) > 200\,\mathrm{GeV}$	1.82×10^{8}	403	
$M_R > 800, R^2 > 0.08$	57557	224	
Trigger	50164	216	67037
no lepton	39666	142	56220
$n_b \geq 1$	12187	119	18164
$n_W \geq 1$	1350	28	1817
S	160	23.4	187
S'	1189	4.55	1630
$n_b = 0$	15691	5.65	20667
$n_{aW} \geq 1$	1923	0.667	2712
Q	1603	0.07	2240
Q'	220	0.566	223
1 lepton	9699	65.0	10008
$n_b \geq 1$	3470	54	3930
$n_W \geq 1$	666	12.3	770
T	151	1.2	153
$n_b = 0$	3329	2.54	3165
$n_Y \geq 1$	786	1.19	581
W	150	0.06	116

7.6.5 Closure tests

In this analysis, we do not explicitly estimate the background in the signal region from the observations in the control regions. Rather, we create a prior distribution for the four background components ($t\bar{t}$+single top, $W(\to \ell\nu)$+jets, multijet, and all others) of the signal regions, that incorporates all statistical and systematic uncertainties, as will be described in detail in Section 7.7. However, in order to verify that the control regions defined in the previous sections provide adequate data-driven models for the backgrounds in the signal region and that the translations between different regions behave as expected, we perform two cross checks, taking into account statistical uncertainties only. The first cross check also illustrates the relations between signal and control regions in a more direct way compared to the full likelihood implementation, where these relations might be less obvious.

First cross check

In the first cross check, we predict the background in a signal-like control region, denoted by S', defined by inverting the $\Delta\phi_{min}$ requirement while preserving the rest of the selection, see Table 7.10. The estimated number of events, $\widehat{N}$, in the S' region for the QCD multijet, $W(\to \ell\nu)$+jets, and $t\bar{t}$+single top (denoted $t\bar{t}+t$) processes is computed from the observations, N_{obs}, in the Q, T, and W regions as follows,

$$\widehat{N}^{S'}_{\text{QCD}} = \left(N^{Q}_{\text{obs}} - N^{Q}_{\text{other,MC}}\right) / \left(\frac{N^{Q}_{\text{QCD}}}{N^{S'}_{\text{QCD}}}\right)_{\text{MC}}, \tag{7.81}$$

$$\widehat{N}^{S'}_{W\ell\nu} = \left(N^{W}_{\text{obs}} - N^{W}_{\text{other,MC}}\right) / \left(\frac{N^{W}_{W\ell\nu}}{N^{S'}_{W\ell\nu}}\right)_{\text{MC}}, \tag{7.82}$$

$$\widehat{N}^{S'}_{t\bar{t}+t} = \left(N^{T}_{\text{obs}} - \widehat{N}^{T}_{\text{QCD}} - N^{T}_{\text{other,MC}}\right) / \left(\frac{N^{T}_{t\bar{t}+t}}{N^{S'}_{t\bar{t}+t}}\right)_{\text{MC}}, \tag{7.83}$$

where the estimated number of multijet events in the T control region is given by,

$$\widehat{N}^{T}_{\text{QCD}} = \left(N^{Q}_{\text{obs}} - N^{Q}_{\text{other,MC}}\right) / \left(\frac{N^{Q}_{\text{QCD}}}{N^{T}_{\text{QCD}}}\right)_{\text{MC}}. \tag{7.84}$$

In these equations, $N_{\text{other,MC}}$ represents the total contribution of all other processes apart from the ones mentioned explicitly, as determined from simulation. Because of the purity of the control regions, these contributions are small. As can be seen from Table 7.12, $N^{T}_{\text{QCD,MC}} = 0$ for the nominal choice of systematic uncertainties. The formulae above can thus be simplified since $\widehat{N}^{T}_{\text{QCD}} = 0$. This is, however, not necessarily the case for other choices of systematic variations. This relation between the T and Q regions is, therefore, still used to constrain the expected multijet background in the T region during the final background estimate.

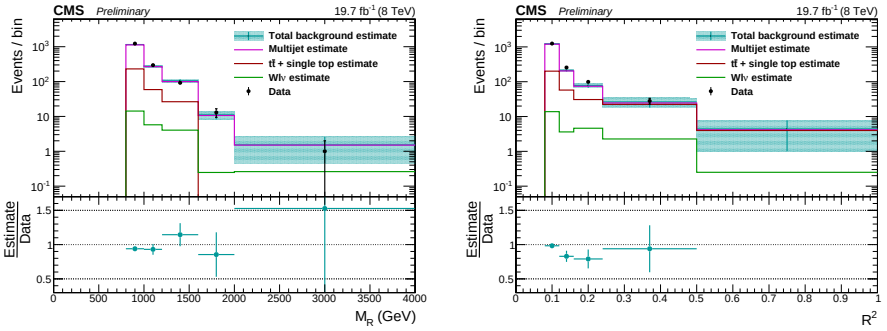

Figure 7.44: Projection of the 2D prediction on the M_R (left) and R^2 (right) axes for the closure test predicting the $\Delta\phi_{min}$ sideband region S'. The uncertainties shown are statistical only and the horizontal error bars only indicate the bin width.

The total estimated background in the S' region is

$$\hat{N}^{S'} = \sum_i \hat{N}_i^{S'}, \tag{7.85}$$

where i runs over all background processes. For the smaller backgrounds, $\hat{N}_i^{S'}$ is determined by simulation. The estimation of backgrounds is done bin-by-bin in the (M_R, R^2) space. However, the estimated scale factors are global because the statistical precision is not sufficient to yield reliable bin-by-bin estimates. The expected global scale factors, which we will denote by κ, are defined in Section 7.7, which also describes how they are calculated from the simulated data.

Figure 7.44 shows the projection on the M_R and R^2 axes of the predicted and observed distributions. The prediction, which includes only statistical uncertainties for this cross check, agrees with observation within 20%. This test of the background modelling shows that it is feasible to estimate a multicomponent background in a signal-like region using the control regions we have defined in Section 7.6.4. In this test, aspects of the modelling in simulation, such as the b tagging, the translation between lepton multiplicities, and certain aspects of the W tagging, have been verified.

Second cross check

In a second cross check, we use the Q region to estimate the background in a more signal-like Q region, denoted by Q', where $\Delta\phi_{min} > 0.5$. The selection is also summarized in Table 7.10. The estimated background in the Q' region, $\hat{N}^{Q'}$, is computed as

$$\hat{N}^{Q'} = N_{\text{obs}}^{Q} \frac{N_{\text{MC}}^{Q'}}{N_{\text{MC}}^{Q}}, \tag{7.86}$$

where N_{obs}^{Q} is the observed data count in the Q region, and N_{MC} includes all contributing simulated background processes.

This test assesses the degree to which the simulated distribution of $\Delta\phi_{min}$ as well as its extrapolation from the Q region, which has $\Delta\phi_{min} < 0.3$, to the S region, with $\Delta\phi_{min} > 0.5$, are reliable. The comparison between prediction and observation is shown in Fig. 7.45. The level of discrepancy, $\sim 40\%$, between the prediction and observation in this cross check is incorporated as a systematic uncertainty in the global scale factors. How this is done technically is described in Section 7.7.

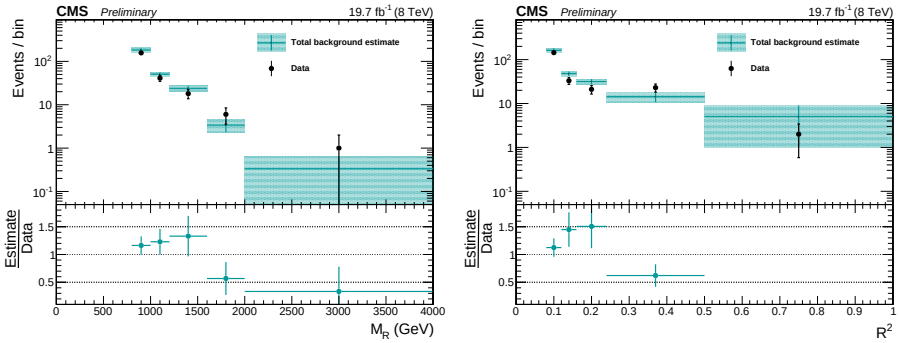

Figure 7.45: Projection of the 2D prediction on the M_R (left) and R^2 (right) axes for the closure test predicting the background in region Q', as defined in the text. The uncertainties shown are statistical only and the horizontal error bars only indicate the bin width.

7.7 Statistical modelling

The razor boost analysis defines one signal region and three control regions, all of which are divided into 25 bins on the (M_R,R^2) space. The control regions are used to constrain the backgrounds in the signal region. The structure of the probability model, which comprises a likelihood for the signal region and an evidence based prior for the background components, is illustrated in Fig. 7.46.

The expected background contributions in the signal region S, represented by the blue circle in the figure, are denoted by b^S_{process}. The index representing the (M_R,R^2) bin has been suppressed in this notation. It is important to remember, however, that the background estimation is done bin-by-bin. The three main background components, b^S_{QCD}, $b^S_{t\bar{t}+t}$, and $b^S_{W\ell\nu}$, are constrained by data in the three control regions, Q, T, and W, respectively. The expected background count b^S_{oth} comprises all the other, smaller, background processes, and is constrained using MC simulated data.

The control regions are shown in the green, red, and magenta circles in Fig. 7.46. The connections between the different circles indicate the relations between the signal and control regions. These relations are encapsulated in the global translation factors $\kappa^{A/B}_{\text{process}}$, where A and B represent any two regions. Consider the W region as example, where we see a link between the W and S regions. This link exists because we use the W region to constrain the

$W(\to \ell\nu)$+jets background in the signal region, $b^S_{W\ell\nu}$. The way this is done is by relating the $W(\to \ell\nu)$+jets contribution in both regions via the scale factor $\kappa^{W/S}_{W\ell\nu}$,

$$b^W_{W\ell\nu} = \kappa^{W/S}_{W\ell\nu} b^S_{W\ell\nu}. \tag{7.87}$$

The total expected count in the W region is just a sum over the components listed within the coloured circle,

$$b^W_{\rm total} = \kappa^{W/S}_{W\ell\nu} b^S_{W\ell\nu} + b^W_{\rm oth}. \tag{7.88}$$

The count for the small other components, $b^W_{\rm oth}$, is again constrained using simulation. The observed data count in the W region will be denoted by N^W. As mentioned several times before, the scale factors κ are global scale factors, meaning that they are not computed separately for each ($\rm M_R$,$\rm R^2$) bin, but rather integrated across all 25 bins. For the example of the W region this becomes,

$$\kappa^{W/S}_{W\ell\nu} = \frac{\sum_{i=1}^{25} b^W_{W\ell\nu,{\rm MC},i}}{\sum_{i=1}^{25} b^S_{W\ell\nu,{\rm MC},i}}, \tag{7.89}$$

where we have added the subscript MC to indicate that MC simulation is used to determine these counts. The total expected counts in the T and Q regions can be computed in a similar fashion, as a sum of the contributions listed in the respective coloured circles in the figure. For the T region there is a small complication because we also use the Q region to constrain the multijet background in the T region.

The statistical analysis of the set of observations, $\{N^S_i\}$, in the signal region is based on a likelihood function, $L(\sigma)$, given by

$$L(\sigma) \equiv \int \left[\prod_{i=1}^{M} p(N^S_i|\sigma,\mathcal{L},\theta_i)\right] \pi(\theta)\,\pi(\mathcal{L})\,d\theta\,d\mathcal{L}, \tag{7.90}$$

where σ is the total signal cross section, and our parameter of interest. The number of bins on the ($\rm M_R$,$\rm R^2$) space is $M = 25$, N^S_i is the observed count in bin i, and the bin-by-bin parameters ϵ, $b^S_{\rm QCD}, b^S_{t\bar{t}+t}, b^S_{W\ell\nu}$, and $b^S_{\rm oth}$ are denoted collectively by θ. The parameter ϵ represents the M signal efficiencies (including acceptance) for a given signal model. The function $\pi(\mathcal{L})$ is the integrated luminosity prior and $\pi(\theta)$ is an evidence based prior constructed from observations in the control regions and the four global scale factors $\kappa^{A/B}_{\rm process}$ determined by simulated data. As mentioned, Fig. 7.46 shows which control regions provide constraints on each of the background parameters, $b^S_{\rm process}$. The likelihood per ($\rm M_R$,$\rm R^2$) bin is taken to be

$$p(N^S|\sigma,\mathcal{L},\theta) = {\rm Poisson}(N^S, \epsilon\sigma\mathcal{L} + b^S_{\rm QCD} + b^S_{t\bar{t}+t} + b^S_{W\ell\nu} + b^S_{\rm oth}). \tag{7.91}$$

The likelihood for all bins is then computed as a product of all per-bin likelihoods. It is clear from the left-hand side of Eq. 7.90 that the final analysis likelihood only depends on the signal cross section. The dependence on all other (nuisance) parameters has been removed through marginalization.

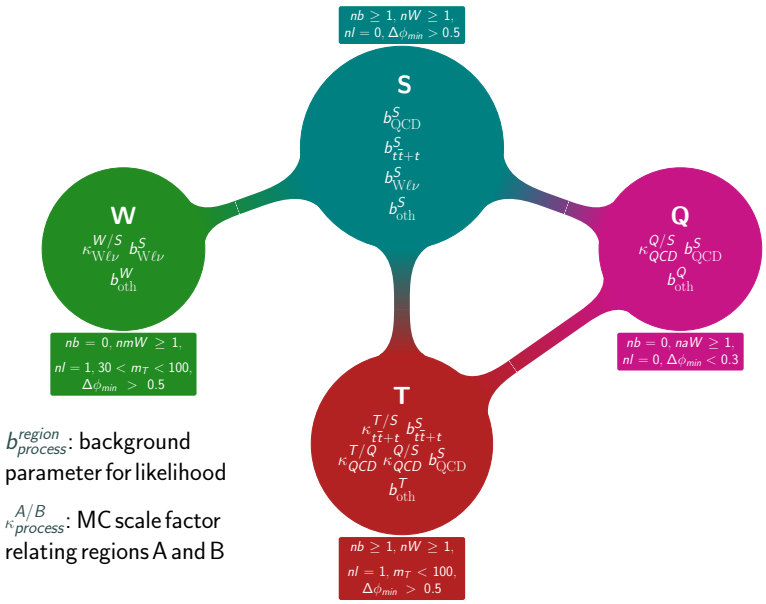

Figure 7.46: Definition of, and relationship between, the signal (S) and control (Q, T, W) regions and their relationship to the bin-by-bin background parameters $b^{\text{region}}_{\text{process}}$ for a given region and background process, as well as the four global scale factors $\kappa^{A/B}_{\text{process}} = \sum_i b^{A}_{\text{process,MC,i}} / \sum_i b^{B}_{\text{process,MC,i}}$, where the sum is over all 25 (M_R,R^2) bins of the simulated data. The total expected background, per bin, is the sum of the terms shown for each region. Furthermore, associated with each bin of each region is an observed count N^{region}, a simulated count $N^{\text{region}}_{\text{process,MC}}$, and a count $N^{\text{region}}_{\text{oth,MC}}$ equal to the sum of the smaller backgrounds, with associated parameter $b^{\text{region}}_{\text{oth}}$.

The integral in Eq. (7.90) is approximated by Monte Carlo integration by sampling the priors $\pi(\mathcal{L})$ and $\pi(\theta)$,

$$L(\sigma) \approx \frac{1}{J}\sum_{j=1}^{J}\prod_{i=1}^{M} p(N_i^S|\sigma,\mathcal{L}_j,\theta_j), \tag{7.92}$$

using J points θ_j randomly sampled from the priors. The priors for the expected integrated luminosity $\mathcal{L}$, signal efficiencies ϵ, and simulated background counts $b^{\text{region}}_{\text{process,MC}}$ (which enter the computation of the κ factors), are modelled with gamma densities,

$$\text{gamma}(x,\gamma,\beta) = \beta^{-1}(x/\beta)^{\gamma-1}\exp(-x/\beta)/\Gamma(\gamma), \tag{7.93}$$

in which the mode is set to c and the variance to δc^2, where $c \pm \delta c$ denotes either the measured integrated luminosity, or for a given bin of a given region and process, the simulated signal efficiency, or the simulated background count. This yields the gamma density parameters,

$$\gamma = [(k+2) + \sqrt{(k+2)^2 - 4}]/2, \tag{7.94}$$

$$\beta = [\sqrt{c^2 + 4\delta c^2} - c]/2, \tag{7.95}$$

where $k = (c/\delta c)^2$. For empty bins, we set $\gamma = 1$ and the bin value is constrained to zero by setting the β parameter to 10^{-4}.

The prior for the expected signal efficiencies and background counts, $\pi(\theta)$, includes the uncertainties coming from the various systematic effects, which will be described in more detail in Section 7.8. The prior is modelled hierarchically,

$$\pi(\theta) = \int \pi(\theta|c)\,\pi(c|\phi)\pi(\phi)\,dcd\phi, \tag{7.96}$$

where c is a simulated count or efficiency in an (M_R,R^2) bin and ϕ represents parameters that characterize the independent sources of systematic uncertainty. The integral in Eq. (7.96) is evaluated via MC integration in the following way.

First, ϕ values are sampled from $\pi(\phi)$, by drawing random numbers from a Gaussian variate. Then, c values are sampled from $\pi(c|\phi)$, followed by θ values from $\pi(\theta|c)$. The sampling from $\pi(\theta|c)$ is straightforward because the functional form is known. For the simulated counts and efficiencies the aforementioned gamma densities are used, while the observed counts in the control regions are modelled with Poisson densities. However, $\pi(c|\phi)$ has an unknown shape, and the sampling of c thus requires running the analysis (in particular the event selection) multiple times, once for each sampling of the systematic uncertainties. The novel feature of the razor boost analysis is that the independent sources of systematic uncertainty are sampled simultaneously, and are then marginalized. This has the very nice consequence that any correlations between systematic uncertainties, processes and regions are properly taken into account. Practically, we produce an ensemble of sets of ($\text{M}_\text{R}, \text{R}^2$) histograms for the simulated backgrounds and efficiencies, for all signals under consideration, which automatically incorporate all statistical dependencies without the need to model them explicitly.

Once the ensemble of sets of (M_R,R^2) histograms is produced, the sampling proceeds as follows:

1. sample the integrated luminosity parameter;
2. sample the efficiency parameters, ϵ, for every considered signal model;
3. sample the parameters $b^{\text{region}}_{\text{process,MC}}$ of the simulated background densities and sum their values over the M bins;
4. compute the κ parameters from the appropriate background sums, for example as in Eq. 7.89;
5. scale $\kappa^{Q/S}_{\text{QCD}}$ by a random Gaussian variate of unit mean and standard deviation of 0.33 to account for additional uncertainty due to deficiencies in the simulated data as determined from the second closure test (Section 7.6.5), resulting in a total uncertainty of 40%;
6. sample the background parameters b^S_{QCD}, $b^S_{t\bar{t}+t}$, and $b^S_{W\ell\nu}$, from the Poisson models of the control regions; for example, for region Q, we map Poisson($N^Q, \kappa^{Q/S}_{\text{QCD}} b^S_{\text{QCD}} + b^Q_{\text{oth}}$) to a posterior density in b^S_{QCD} using Bayes theorem and a flat prior and sample b^S_{QCD} from that density.

The expected background counts for each signal bin are then obtained from the sampled background parameters, and are compared to the observations in data. The final results are shown in Section 7.9.

In the absence of a signal, we determine limits on the total signal cross section using the CLs criterion [182] with the test statistic $t_\sigma = 2\ln[L(\hat{\sigma})/L(\sigma)]$ when $0 \leq \hat{\sigma} \leq \sigma$, and $t_\sigma = 0$ when $\hat{\sigma} > \sigma$. Large values of t_σ indicate incompatibility between the best fit hypothesis $\sigma' = \hat{\sigma}$ and the entertained hypothesis $\sigma' = \sigma$. We calculate the p-values

$$p_0 = \text{Prob}(t_\sigma > t_{\sigma,obs} | \sigma' = 0), \tag{7.97}$$

and

$$p_\sigma = \text{Prob}(t_\sigma > t_{\sigma,obs} | \sigma' = \sigma), \tag{7.98}$$

needed to calculate CLs(σ) $= p_\sigma/p_0$, by simulation. The quantity $t_{\sigma,obs}$ denotes the observed values of the test statistic, one for each hypothesis $\sigma' = \sigma$.

7.8 Systematic uncertainties

The input to the statistical analysis is an ensemble of histograms in the (M_R, R^2) plane that incorporates systematic uncertainties in the simulated signal and background samples. The independent systematic effects, described below, are sampled simultaneously. This is one of the characteristics that sets this analysis apart. For each sampled systematic effect, the same zero mean, unit variance, Gaussian variate is used in the calculation of the random shift of the systematic effect for all the signal and background models. Likewise, the same randomly

sampled parton distribution functions are used for each event and for all signal and background models. In this way, the statistical dependencies among all bins of the signal and background models are correctly, and automatically, modelled. The sampling of the systematic effects is repeated several hundred (N) times. In all cases, except for the PDFs, the systematic uncertainties are in the scale factors applied to the simulated samples to correct them for modelling deficiencies. In the next subsections I will discuss each source of systematic uncertainty in more detail.

7.8.1 Jet energy scale corrections

Jet energy scale corrections (JEC) map the measured jet energy deposition in the detector back to the particle level. Within CMS the JEC are applied in sequential levels, each correcting for a different effect. Details on the various levels were given in Section 6.4.2. For our discussion here, the important thing to remember is that each level of correction is simply a scaling of the four-momentum of the jet by a scale factor that depends on the jet p_{T} and η.

Uncertainties on JEC originate from various uncorrelated sources. Given M such sources, the full correction on the p_{T} of a jet, according to a random shift within the uncertainties, becomes

$$s(p_{\mathrm{T}}, \eta, \alpha_i) = \sum_{i=1}^{M} \alpha_i S_i(p_{\mathrm{T}}, \eta) \tag{7.99}$$

where $S_i(p_{\mathrm{T}}, \eta)$ is a p_{T} and η-dependent JEC uncertainty for a source i and α_i are weights randomly sampled from a Gaussian distribution with zero mean and unit width. Here, the weights α_i are different for each source, but are universal for each jet and each event when considering a given systematic sampling. Dealing with the myriad of uncertainty sources is not so straightforward in practice, and also not necessary for a new physics search. As recommended by the POG providing the jet energy scale corrections and uncertainties, we will therefore take on a much simpler approach, using a total uncertainty $S(p_{\mathrm{T}}, \eta)$ and a random number α from a single Gaussian,

$$s(p_{\mathrm{T}}, \eta, \alpha) = \alpha S(p_{\mathrm{T}}, \eta) = \alpha \sqrt{\sum_{i}^{M} s_i^2(p_{\mathrm{T}}, \eta)}. \tag{7.100}$$

The jet p_T for a given random variation becomes:

$$p_{\mathrm{T}}^{\mathrm{corr}} = (1 + s(p_{\mathrm{T}}, \eta, \alpha))\, p_{\mathrm{T}}^{\mathrm{orig}}. \tag{7.101}$$

This calculation is repeated N times, each time using a different number sampled from the Gaussian. The overall effect of the JEC uncertainty on a given yield is obtained from the distribution of resulting N yields. As was done for the jet energy scale corrections themselves, the associated uncertainties are also propagated to the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$.

7.8.2 Parton distribution functions

In order to evaluate the systematic uncertainty arising from imperfect knowledge of the parton distribution functions, we use three PDF sets, `CT10` [176], `MSTW2008lo68cl` [183], and `NNPDF23_lo_as_0130_qed` [184], which are recommended by the PDF4LHC group [185, 186]. Since we take into account full correlations within systematic variations, we need a way to sample randomly also from the PDF uncertainties, such that the PDF variations can be incorporated directly in our workflow. Of the three recommended PDF sets, `NNPDF` already presents the PDF eigenvectors as a randomly distributed set, while the other groups provide eigenvectors obtained by varying the PDF fit parameters by ± 1 standard deviation. However, the recently developed `LHAPDF6` [187] offers a formal way to convert the latter sets into the randomly distributed sets we need. We have used the program `hessian2replicas` to generate randomly distributed PDF sets with 100 members each, for both `CT10` and `MSTW2008lo68cl`.

Given a sampled set i, for PDF set K, and the PDF set O with which the events were simulated, events are reweighted using the scale factors,

$$\mathrm{SF}_{K,i} = \frac{w_{K,i}}{w_O}, \tag{7.102}$$

where the weights w are products of the event-by-event PDFs for both colliding partons. The PDF set O was listed for each sample in Section 7.5.2. The overall PDF uncertainty on the background counts or signal efficiencies is derived from the distribution of $\mathrm{SF}_{K,i}$ obtained by a random selection of K and i.

To study the effect of the different PDF sets on the simulated counts used in the background estimation, we compute the distribution of simulated counts corresponding to a random selection of PDF members from the three considered PDF sets. A Gaussian distribution is then fitted to these distributions. Its width indicates the uncertainty induced by the given PDF set. Figures 7.47-7.51 show the result of this procedure for the simulated background counts that are used to compute the translation factors κ. There is a Gaussian distribution for each PDF set, and the arrow indicates the value for a given count as obtained using the PDF set that was used during the generation of the MC samples. From these figures we conclude that the overall systematic uncertainty resulting from the parton distribution functions on the background counts is due to both the difference in the nominal values of the PDF sets, and the spread within the separate PDF sets.

Figure 7.52 shows the equivalent for an example signal point. Here we see that the spread of the `CT10` PDF dominates the total uncertainty on the signal efficiency. We also note that we have computed the signal efficiencies with respect to the nominal PDF set member of the considered PDF rather than the original PDF used for the generation of the samples. This is done to make sure that only the PDF effects on the acceptance are included in the experimental uncertainties. The PDF effect on the expected signal cross section is taken into account separately.

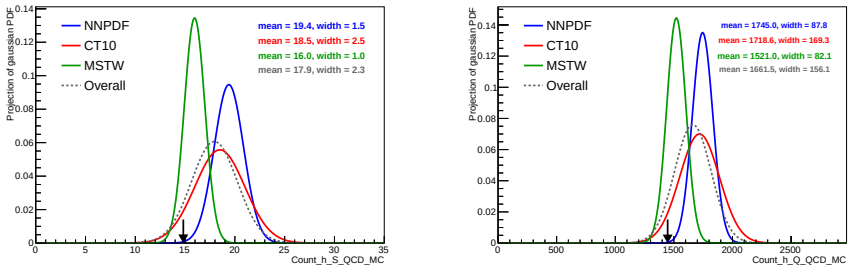

Figure 7.47: Influence of different PDF sets on the MC counts entering $\kappa_{\mathrm{QCD}}^{Q/S}$.

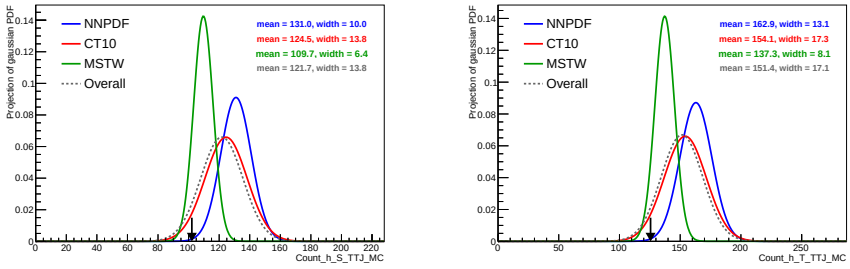

Figure 7.48: Influence of different PDF sets on the MC counts entering $\kappa_{t\bar{t}+t}^{T/S}$.

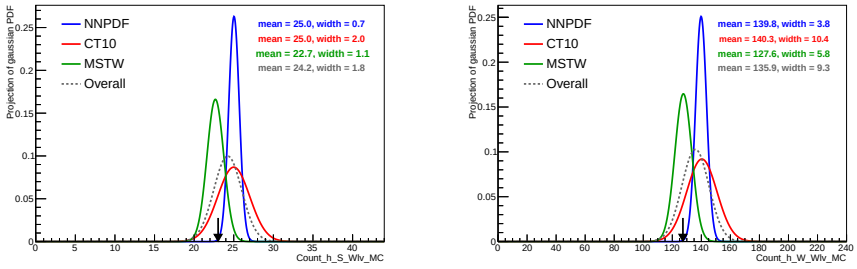

Figure 7.49: Influence of different PDF sets on the MC counts entering $\kappa_{W\ell\nu}^{W/S}$.

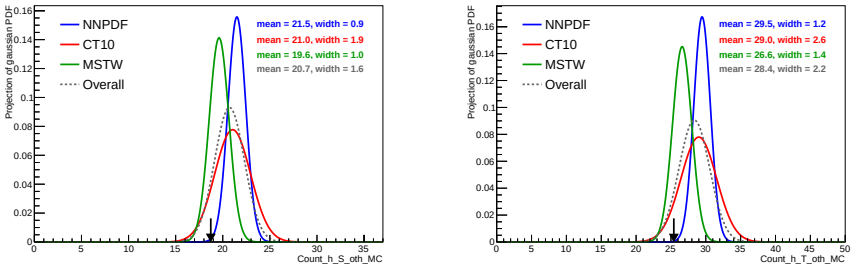

Figure 7.50: Influence of different PDF sets on the MC counts in the S (left) and T (right) region for the backgrounds that are taken directly from the simulation.

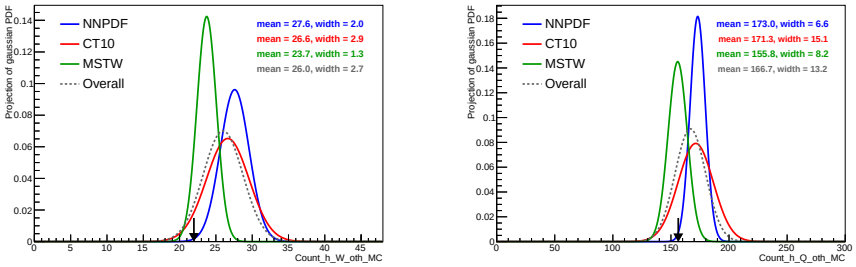

Figure 7.51: Influence of different PDF sets on the MC counts in the W (left) and Q (right) region for the backgrounds that are taken directly from the simulation.

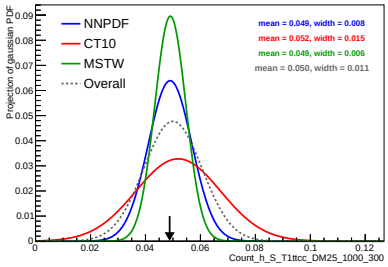

Figure 7.52: Influence of different PDF sets on the signal efficiency for the T1ttcc signal point with $m_{\tilde{g}} = 1000\,\text{GeV}, m_{\tilde{t}_1} = 325\,\text{GeV}, m_{\tilde{\chi}_1^0} = 300\,\text{GeV}$.

7.8.3 Trigger efficiency

Simulated events are weighted by a trigger efficiency, ϵ_{trig}, according to their values of H_{T} and first jet p_{T} (see Section 7.5.1). These trigger weights have an associated uncertainty coming from the statistical precision of the data samples used to perform the measurement, and the effect of changing the analysis level selection used to measure the efficiency. The trigger efficiency uncertainty in each bin, as a function of H_{T} and leading jet p_{T}, is taken to be the maximum of the statistical uncertainty in the efficiency after imposing the baseline selection and the difference between the efficiencies before and after applying the baseline selection. Magnitudes of the plus and minus uncertainties, $\delta\epsilon^{+}_{\text{trig}}$ and $\delta\epsilon^{-}_{\text{trig}}$, were shown in Fig. 7.23. The event weight w_{trig} to be applied to the simulated events corresponding to a given systematic sampling is computed based on the efficiency and uncertainty as follows:

$$w_{\text{trig}} = \epsilon_{\text{trig}}(H_{\text{T}}, p_{\text{T}}^{\text{j}_1}) + \sigma_{\text{trig}}\delta\epsilon^{+}_{\text{trig}}(H_{\text{T}}, p_{\text{T}}^{\text{j}_1}), \text{ if } \sigma_{\text{trig}} > 0 \quad (7.103)$$

$$w_{\text{trig}} = \epsilon_{\text{trig}}(H_{\text{T}}, p_{\text{T}}^{\text{j}_1}) + \sigma_{\text{trig}}\delta\epsilon^{-}_{\text{trig}}(H_{\text{T}}, p_{\text{T}}^{\text{j}_1}), \text{ if } \sigma_{\text{trig}} < 0 \quad (7.104)$$

where σ_{trig} is the random Gaussian number associated to that particular systematic sampling.

7.8.4 b tagging

The b tagging performance differs between data and simulation, and differs further between CMS FullSim and FastSim. The simulated events are therefore corrected by applying data/FullSim and FullSim/FastSim b tag efficiency scale factors, which depend on the flavour, p_{T}, and η of the jets in the event.

There are several methods available to reweight events using the b tag SFs provided by the CMS BTAG group [188]. Here we choose a method where we consider the number of b tagged jets in an event to be fixed to what is given by simulation. For a given event, the b-tagging efficiencies for truth b, c and udsg jets, corresponding to the relevant tagging algorithm and working point, are used to compute the probability that the event has the given b tagged jet multiplicity. We compute $P(\text{sim})$ using efficiencies, ϵ^{sim}, obtained from simulated events, and we obtain $P(\text{data})$ using efficiencies, ϵ^{data}, that represent data,

$$P(\text{sim}) = \prod_i \epsilon_i^{\text{sim}} \prod_j (1 - \epsilon_j^{\text{sim}}) \quad (7.105)$$

$$P(\text{data}) = \prod_i \epsilon_i^{\text{data}} \prod_j (1 - \epsilon_j^{\text{data}}) \quad (7.106)$$

where

$$\epsilon^{\text{data}} = \text{SF}^{\text{data}}_{\text{sim}} \epsilon^{\text{sim}}, \quad (7.107)$$

i runs over the jets that are tagged by the considered b tagging algorithm, and j runs over the jets that are not tagged. The SM samples are simulated with FullSim, so we scale the efficiency with $\text{SF}^{\text{data}}_{\text{sim}} = \text{SF}^{\text{data}}_{\text{Full}}$. However, the signal samples have been generated using FastSim, so the overall scale factor for those

samples becomes $\mathrm{SF}_{\mathrm{sim}}^{\mathrm{data}} = \mathrm{SF}_{\mathrm{Full}}^{\mathrm{data}} \cdot \mathrm{SF}_{\mathrm{Fast}}^{\mathrm{Full}}$. The ratio of the two probabilities gives us the weight with which to scale the event,

$$w_{\mathrm{btag}} = \frac{P(\mathrm{data})}{P(\mathrm{sim})}. \tag{7.108}$$

The b tag scale factors have associated uncertainties, δSF, which are the source of the b tag systematic uncertainty on the background prediction. These uncertainties are taken into account in the following way. N numbers σ_n are sampled from a Gaussian with zero mean and unit variance (different σ's are sampled for FullSim and FastSim, and also for different jet truth flavours). The analysis is then run N times. For each run n we use the corresponding σ_n to compute ϵ^{data} for the FullSim or FastSim samples as follows,

$$\text{FullSim: } \epsilon^{\mathrm{data}} = \left(\mathrm{SF}_{\mathrm{Full}}^{\mathrm{data}} + \sigma_n \, \delta\mathrm{SF}_{\mathrm{Full}}^{\mathrm{data}}\right) \cdot \epsilon^{\mathrm{Full}}, \tag{7.109}$$

$$\text{FastSim: } \epsilon^{\mathrm{data}} = \left(\mathrm{SF}_{\mathrm{Full}}^{\mathrm{data}} + \sigma_n \, \delta\mathrm{SF}_{\mathrm{Full}}^{\mathrm{data}}\right) \cdot \left(\mathrm{SF}_{\mathrm{Fast}}^{\mathrm{Full}} + \sigma'_n \, \delta\mathrm{SF}_{\mathrm{Fast}}^{\mathrm{Full}}\right) \cdot \epsilon^{\mathrm{Fast}}. \tag{7.110}$$

For each of the N runs, ϵ^{data} will be different, and will lead to a different event weight w_{btag} such that, overall, the weighted event yields for a given simulated sample will differ. We obtain the systematic uncertainty from the ensemble of the resulting N yields.

In this analysis, we use the CSVM and CSVL b taggers, see Section 6.4.2, which are always used independently of each other. Therefore, for a given event, we compute separate weights w_{btag} using CSVM and CSVL, and use them only in the relevant regions. The weight obtained with the CSVM tagger is used in the S and T regions, and the CSVL weight in the W and Q regions.

7.8.5 W tagging

The W boson tag efficiency, and the fake rate for W boson tag, W boson mass-tag, and W boson anti-tag differ between data and simulation, as well as between FullSim and FastSim. Data/FullSim and FullSim/FastSim scale factors, whose uncertainties depend on the jet p_{T}, are applied to the simulated samples to correct for this mismodelling. Some of these scale factors were derived specifically for this analysis and were described in detail in Section 7.4. The uncertainties on these scale factors are used in the same way as for the other sources of systematic uncertainty. For each systematic sampling, the scale factors are varied according to the sampled uncertainty and propagated to the corresponding event weights in the following way

$$w_{\mathrm{Wtag}} = \mathrm{SF}_{\mathrm{Wtag}}(p_{\mathrm{T}}) + \sigma_{\mathrm{Wtag}} \delta\mathrm{SF}_{\mathrm{Wtag}}(p_{\mathrm{T}}), \tag{7.111}$$

7.8.6 Lepton identification

The razor boost analysis requires the presence of a single loose electron or muon in the definition of T and W regions. We therefore apply lepton scale factors

to our simulated events to correct for mismodelling of the lepton identification. The uncertainties on these scale factors are incorporated as a systematic uncertainty on the background prediction.

For electrons, we use p_{T} and η-dependent scale factors and associated uncertainties, as derived by the CMS EGamma POG [189]. The overall event weight associated with a given systematic variation of the electron scale factor, is given as:

$$w_{\mathrm{e}} = \mathrm{SF}_{\mathrm{e}}(p_{\mathrm{T}}, \eta) + \sigma_{\mathrm{e}} \delta \mathrm{SF}_{\mathrm{e}}(p_{\mathrm{T}}, \eta), \tag{7.112}$$

where σ_{e} is the random Gaussian number associated with the given systematic variation. The scale factors for muons are approximately equal to unity, with negligible uncertainties. Their variation is, consequently, not taken into account.

7.8.7 Initial State Radiation

Deficiencies in the modelling of initial state radiation are corrected by reweighting the signal samples using an event weight that depends on the p_{T} of the recoiling system, which in our case is the $\tilde{g}\tilde{g}$ system. More information on this reweighting and how it was derived, was given in Section 6.4.4. The associated uncertainty on the ISR reweighting (see Table 6.12 for the numerical values) is incorporated as a systematic uncertainty on the signal efficiency.

7.8.8 Top quark p_{T} spectrum

Differential top-quark-pair cross section analyses have shown that the shape of the p_{T} spectrum of top quarks in data is softer than predicted by simulation. To account for this, we reweight events based on the p_{T} of the generator level t and $\bar{t}$ quarks in the $t\bar{t}$ simulation, as explained in Section 6.4.4. The uncertainty associated with this reweighting is equal to the full size of the reweighting, and is propagated as a systematic uncertainty to the final background estimation.

7.8.9 Pileup

Simulated events are reweighted such that their pileup distribution matches the observed pileup distribution, see Section 6.4.4. The uncertainty on this procedure is largely driven by the uncertainty on the minbias cross section. We consider a variation in the minbias cross section of $\pm 5\%$, thereby changing the shape of observed pileup distribution and therefore the resulting pileup event weights. The difference in weights is taken as a measure of the uncertainty in the pileup distribution, and is multiplied by the random number, σ_{PU}, associated to each given sampling. The pileup weight that is applied in each sampling is given by

$$w_{\mathrm{PU}} = w_{\mathrm{PU}}^{\mathrm{nom}} + \sigma_{\mathrm{PU}} \delta w_{\mathrm{PU}}^{+}, \text{ if } \sigma_{\mathrm{PU}} > 0, \tag{7.113}$$

$$w_{\mathrm{PU}} = w_{\mathrm{PU}}^{\mathrm{nom}} + \sigma_{\mathrm{PU}} \delta w_{\mathrm{PU}}^{-}, \text{ if } \sigma_{\mathrm{PU}} < 0. \tag{7.114}$$

with $w_{\mathrm{PU}}^{\mathrm{nom}}$ the nominal pileup weight, $\delta w_{\mathrm{PU}}^{+} = w_{\mathrm{PU}}^{+5\%} - w_{\mathrm{PU}}^{\mathrm{nom}}$, and $\delta w_{\mathrm{PU}}^{-} = w_{\mathrm{PU}}^{\mathrm{nom}} - w_{\mathrm{PU}}^{-5\%}$.

7.8.10 QCD spectrum

The closure tests described in Section 7.6.5 showed that there is a 40% uncertainty in the QCD multijet scale factor $\kappa_{\mathrm{QCD}}^{Q/S}$ between the signal and Q region. A quantitative verification of this number is given in the next paragraph. The uncertainty is accounted for in the background prediction method by including an additional 33% uncertainty directly on the $\kappa_{\mathrm{QCD}}^{Q/S}$ parameter, as was mentioned in Section 7.7. This is the only systematic uncertainty treated in this way. Every other systematic uncertainty is dealt with during the systematic sampling from which the expected simulated counts are computed.

The level of closure in the second cross check, see Section 7.6.5, can be computed using the ratio between observed data and prediction. We model each bin of that ratio, on the two-dimensional ($\mathrm{M_R}$, R^2) plane, as a Gaussian distribution, with the uncertainty on the ratio for that bin as the width of the Gaussian. We scale each distribution to the number of observed events, thus giving more weight to bins with higher precision, and finally sum over all of them. From this total distribution we then compute the interval around unity that contains approximately 68% of the integral. This procedure is illustrated in Fig. 7.53. On the left the ratio data/prediction with associated statistical uncertainty is shown for each bin in the two-dimensional razor space. On the right-hand side we show the Gaussian distributions and the corresponding 68% interval that constitutes the total systematic uncertainty, 40%, that needs to be applied.

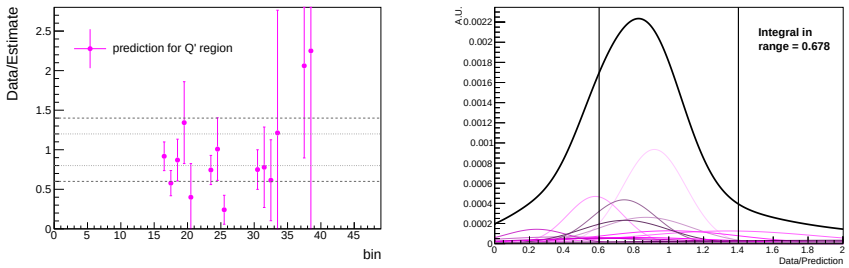

Figure 7.53: [left] Data/prediction for each of the 25 bins in the 2D ($\mathrm{M_R}, \mathrm{R}^2$) plane for the closure test predicting the background in region Q'. Uncertainties are statistical only. [right] We represent the agreement between data and prediction for the closure test predicting the background in region Q' as a Gaussian probability density function for each bin in the 2D ($\mathrm{M_R}, \mathrm{R}^2$) plane. Each bin is shown as a Gaussian in a different shade of magenta. The sum of all Gaussians is depicted in black. Each separate component has been normalized to the weight it carries in the sum.

7.8.11 $\mathrm{Z}(\to \nu\bar{\nu})$+jets in association with heavy flavour

About 8% of the background in the signal region is composed of $\mathrm{Z}(\to \nu\bar{\nu})$+jets events. Since we require the presence of at least one b tagged jet, and given the

known deficiency in modelling Z production in association with heavy flavour jets, we include an extra systematic uncertainty in the $Z(\to \nu\bar{\nu})$+jets contribution.

This uncertainty is estimated using a data control region which is enriched in $Z(\to \ell\bar{\ell})$+jets events. The events are required to satisfy the baseline selection requirements, and to contain exactly two tight leptons (e or μ), of same flavour and opposite sign, with dilepton invariant mass consistent with the Z boson mass, $60 < m_{\ell\bar{\ell}} < 120\,\text{GeV}$. We also require the presence of at least one b tagged jet, and at least one W boson mass-tagged jet. A comparison between data and simulation in this Z-enriched control region is presented in Fig. 7.54. The simulation is seen to overpredict the data.

We estimate the systematic uncertainty in the $Z(\to \nu\bar{\nu})$+jets contribution by first computing bin-by-bin data/simulation ratios, on the (M_R, R^2) space, in this control region. Then, we take the statistical uncertainty in the ratio for each bin as the standard deviation of a Gaussian distribution, normalized to the number of events in that bin. Finally, the Gaussians from all bins are superposed, and the total uncertainty is taken to be the magnitude of the 68% band around a ratio of unity. Figure 7.55 illustrates this procedure, showing the bin-by-bin ratios and the superposed Gaussians. Based on these results, we decide to put an additional uncertainty of 50% on the contribution of $Z(\to \nu\bar{\nu})$ in association with heavy flavour.

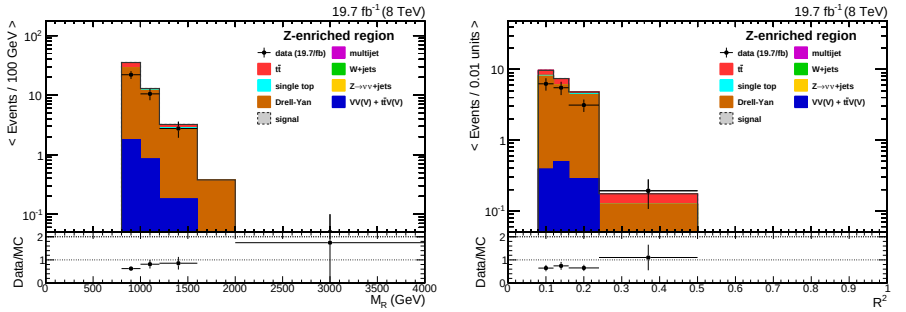

Figure 7.54: Comparison of data and simulation for the M_R (left) and R^2 (right) distributions in the $Z \to \ell\bar{\ell}$ control region with at least one b tagged jet and at least one mass-tagged W boson candidate.

7.8.12 Summary of separate systematic effects

As noted before, all systematic effects are varied simultaneously. However, to assess the effect of each systematic uncertainty individually, each systematic effect i is varied by one standard deviation up and down. The effect on the background and signal processes in the signal region is shown in Table 7.14. The signal values are obtained from averaging over all mass points in the T1ttcc ($\Delta m = 25\,\text{GeV}$) plane. The size of the systematic effects for each separate mass point is shown in the figures in Appendix B. The PDF systematic uncertainties are obtained by running over 100 different PDF set members, sampled from

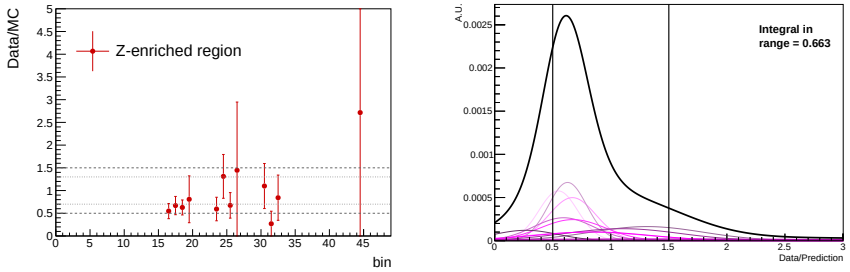

Figure 7.55: [left] Ratio data/simulation for each bin in the 2D (M_R, R^2) plane for the Z-enriched control region mentioned in the text. Uncertainties are statistical only. [right] We represent the agreement between data and simulation for the Z-enriched control region as a Gaussian probability density function for each bin in the 2D (M_R, R^2) plane. Each bin is shown as a Gaussian in a different shade of magenta. The sum of all Gaussians is depicted in black. Each separate component is normalized to the weight it carries in the sum.

all three considered PDF sets, fitting a Gaussian to the efficiency distribution and taking the width of that Gaussian. The last line in the table corresponds to the full sampling of the systematic uncertainties, as used in the background prediction. To obtain this value we again fit a Gaussian to the efficiency distribution obtained from the full systematic sampling including 500 variations. We note that, although the effects of some of these systematic uncertainties on the backgrounds are large, these do not influence our results greatly because only the ratios of simulated background counts enter the statistical analysis, via the κ factors, and not the distributions themselves. Therefore, most of the systematic effects are largely reduced in magnitude.

The dominant systematic uncertainty arises from the parton distribution functions, with an effect of 15-25% on the signal efficiencies and around 10% on the simulated background counts. For the signal samples with a very compressed mass spectrum the uncertainty on the ISR modelling reaches about 20%, compared to only 4-7% for non-compressed spectra. The statistical precision of the control regions is the leading uncertainty for the search bins at large M_R or R^2.

7.9 Results

We present the results of the background prediction for each of the 25 considered search bins in the (M_R, R^2) plane in Fig. 7.56 and Table 7.15. The results are presented as the mean and standard deviation as determined from the sampled prior $\pi(\theta)$ described in Section 7.8. The observations are found to be in good agreement with the Standard Model prediction. No evidence of a signal is found. We do note that for small counts the mean gives a substantially larger value for the background prediction than the mode. This is further illustrated in Fig. 7.57 for two of the predicted background counts. The many bins in the

Table 7.14: Summary of $\pm 1\sigma$ systematic uncertainties for the average signal count of all T1ttcc ($\Delta m = 25\,\text{GeV}$) signal points, and for the total background count in the signal region, unless indicated otherwise, as determined from simulation.

Systematic Effect	Signal	Background
JEC	$+2.2\% - 2.1\%$	$+10.9\% - 5.2\%$
Trigger	$+1.1\% - 3.3\%$	$+3.4\% - 5.7\%$
b tag FullSim	$+2.1\% - 2.3\%$	$+3.9\% - 4.0\%$
b tag FastSim	$+1.2\% - 1.3\%$	-
W tag efficiency Fullsim	$+9.0\% - 8.9\%$	$+4.6\% - 4.6\%$
W tag efficiency FastSim	$+2.2\% - 2.2\%$	-
W tag fake rate FullSim	-	$+1.4\% - 1.4\%$
W anti-tag fake rate FullSim (Q region only)	-	$+2.6\% - 2.6\%$
W mass-tag fake rate FullSim (W region only)	-	$+2.3\% - 2.3\%$
Electron ID (T and W region only)	-	$+0.2\% - 0.2\%$
Pileup	$+0.5\% - 0.5\%$	$+1.0\% - 1.1\%$
ISR	$+6.6\% - 6.6\%$	-
Top p_T spectrum	-	-14.4% 20.5%
$\text{Z} \to \nu\nu +$ heavy flavour	-	$+4.0\% - 4.0\%$
PDF	20.7%	10.7%
All	24.4%	22.1%

high M_R, high R^2 region for which the prediction is around one event, without any observed data are thus no cause for concern. The full distribution of the counts is used for the interpretation of the results which ensures the proper treatment of the non-Gaussian shape.

7.10 Interpretation in terms of simplified model spectra

We interpret our results in terms of the simplified models T1ttcc and T1t1t, whose diagrams were shown in Fig. 7.2. The general characteristics of simplified models were discussed in Section 4.7. These particular models have three free mass parameters: the masses of the gluino, top squark and LSP. The gluino mass is varied between 600 and 1300 GeV, and the LSP mass between 1 and 500 GeV. The mass difference between top squark and LSP, Δm, is kept fixed at 10, 25 or 80 GeV for the T1ttcc model, and at 175 GeV for the T1t1t model.

To get a sense of the expected signal sensitivity, we show the signal efficiencies for the T1ttcc and T1t1t simplified models in Fig. 7.59, and the expected signal counts in Fig. 7.60. Efficiencies of up to 6% in the most highly boosted regimes are reached.

For the T1ttcc model a sharp drop in efficiency is observed for the strip with the lowest considered neutralino mass, $m_{\tilde{\chi}_1^0} = 1\,\text{GeV}$, compared to neighbouring mass points with $m_{\tilde{\chi}_1^0} = 50\,\text{GeV}$. This can be explained by Lorentz boosts. The top squark decays into a charm quark and the LSP. The momentum of the decay

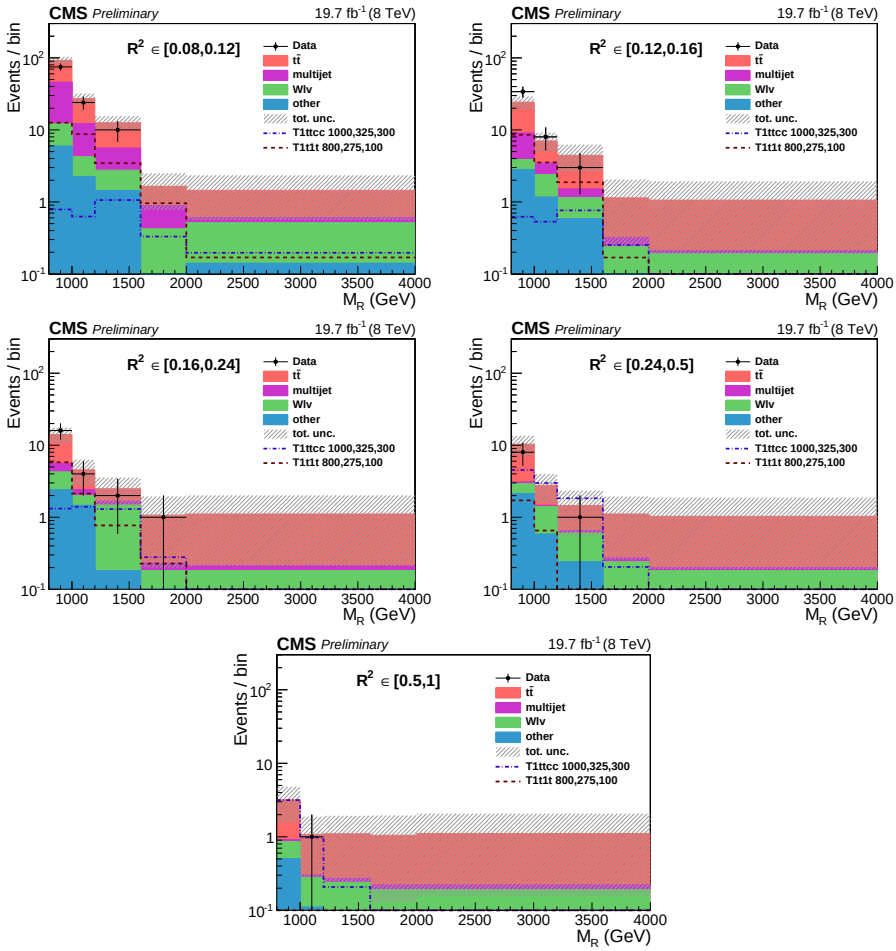

Figure 7.56: Results of the background prediction and comparison with data. Results are shown in bins of M_R for each R^2 strip. The hatched band represents the total uncertainty on the background prediction. Overlaid are two signal distributions corresponding to the T1ttcc model point with $m_{\tilde{g}} = 1\,\text{TeV}$, $m_{\tilde{t}_1} = 325\,\text{GeV}$ and $m_{\tilde{\chi}_1^0} = 300\,\text{GeV}$, and the T1t1t model point with $m_{\tilde{g}} = 800\,\text{GeV}$, $m_{\tilde{t}_1} = 275\,\text{GeV}$ and $m_{\tilde{\chi}_1^0} = 100\,\text{GeV}$.

Table 7.15: Background prediction results and observation in data for all search bins. Uncertainties on the prediction are the combined statistical and systematic uncertainties as obtained from the sampling procedure.

R^2	M_R	$t\bar{t}$	Multijet	$W \to l\nu$	Other	Total	Observed
[0.08,0.12]	[800,1000]	46.7 ± 7.9	33.6 ± 7.6	6.1 ± 1.7	5.9 ± 2.2	92.3 ± 11.3	75
	[1000,1200]	15.0 ± 4.1	8.0 ± 1.9	2.0 ± 0.9	2.2 ± 0.8	27.2 ± 4.7	24
	[1200,1600]	7.0 ± 2.6	2.8 ± 0.7	1.3 ± 0.7	1.4 ± 0.7	12.6 ± 3.0	10
	[1600,2000]	0.8 ± 0.8	0.5 ± 0.2	0.4 ± 0.3	0.1 ± 0.0	1.6 ± 0.9	0
	[2000,4000]	0.8 ± 0.9	0.1 ± 0.1	0.4 ± 0.3	0.1 ± 0.1	1.4 ± 0.9	0
[0.12,0.16]	[800,1000]	15.3 ± 4.5	5.1 ± 1.2	1.1 ± 0.8	2.8 ± 1.1	24.3 ± 4.8	34
	[1000,1200]	3.6 ± 2.0	1.0 ± 0.3	1.2 ± 0.6	1.2 ± 0.6	7.0 ± 2.1	8
	[1200,1600]	2.9 ± 1.7	0.4 ± 0.1	0.6 ± 0.3	0.6 ± 0.4	4.4 ± 1.8	3
	[1600,2000]	0.8 ± 0.9	0.1 ± 0.1	0.2 ± 0.2	0.1 ± 0.0	1.1 ± 0.9	0
	[2000,4000]	0.8 ± 0.8	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	1.1 ± 0.9	0
[0.16,0.24]	[800,1000]	8.5 ± 3.2	1.4 ± 0.4	1.8 ± 0.8	2.4 ± 1.1	14.1 ± 3.5	16
	[1000,1200]	2.2 ± 1.6	0.4 ± 0.2	0.5 ± 0.3	1.5 ± 0.7	4.5 ± 1.8	4
	[1200,1600]	0.8 ± 0.9	0.2 ± 0.1	1.3 ± 0.6	0.2 ± 0.1	2.5 ± 1.1	2
	[1600,2000]	0.8 ± 0.9	0.1 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	1.1 ± 0.9	1
	[2000,4000]	0.9 ± 0.9	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	1.1 ± 0.9	0
[0.24,0.5]	[800,1000]	7.3 ± 3.0	0.1 ± 0.1	0.8 ± 0.5	2.1 ± 1.0	10.3 ± 3.2	8
	[1000,1200]	1.3 ± 1.1	0.1 ± 0.0	0.8 ± 0.4	0.6 ± 0.4	2.8 ± 1.2	0
	[1200,1600]	0.8 ± 0.9	0.1 ± 0.0	0.4 ± 0.2	0.2 ± 0.1	1.4 ± 0.9	1
	[1600,2000]	0.8 ± 0.8	0.0 ± 0.0	0.2 ± 0.2	0.1 ± 0.0	1.1 ± 0.8	0
	[2000,4000]	0.8 ± 0.8	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	1.0 ± 0.8	0
[0.5,1]	[800,1000]	2.3 ± 1.6	0.1 ± 0.1	0.4 ± 0.3	0.5 ± 0.3	3.2 ± 1.6	0
	[1000,1200]	0.8 ± 0.8	0.0 ± 0.0	0.2 ± 0.2	0.1 ± 0.1	1.1 ± 0.8	1
	[1200,1600]	0.8 ± 0.8	0.0 ± 0.0	0.2 ± 0.2	0.1 ± 0.1	1.1 ± 0.8	0
	[1600,2000]	0.8 ± 0.9	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	1.0 ± 0.9	0
	[2000,4000]	0.9 ± 0.9	0.0 ± 0.0	0.2 ± 0.2	0.0 ± 0.0	1.1 ± 0.9	0

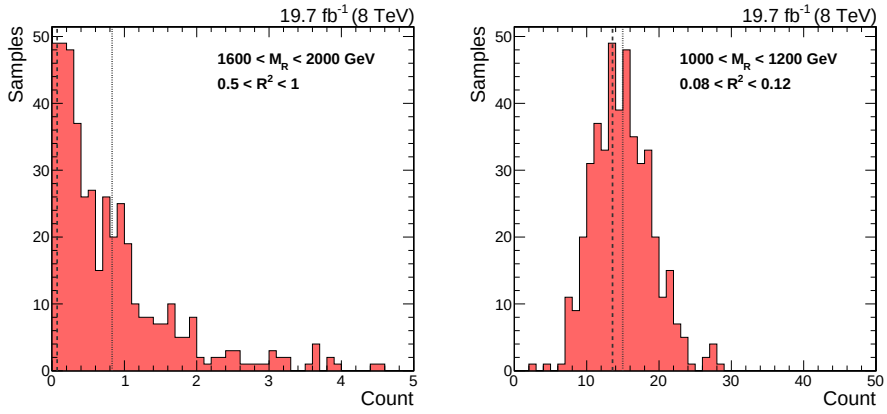

Figure 7.57: [left] Distribution of the expected count for the $t\bar{t}$ and single top background for the search bin defined by $1600 < \mathrm{M_R} < 2000\,\mathrm{GeV}$ and $0.5 < \mathrm{R}^2 < 1$. It is clear that this distribution is highly non-Gaussian with the mode, shown as the dashed line, being much smaller than the mean, shown as the dotted line. [right] The equivalent for the search bin defined by $1000 < \mathrm{M_R} < 1200\,\mathrm{GeV}$ and $0.08 < \mathrm{R}^2 < 0.12$. Here the distribution has a much more Gaussian-like shape, and a smaller difference between mode and mean.

products in the lab frame, p, is related to the momentum and energy in the top squark rest frame, p' and E', by $p = \gamma(p' + \beta E')$. Both decay products have the same momentum in the top squark rest frame, but their energies will depend on their masses. For LSP masses higher than the mass of the charm quark, the LSP will have a higher energy in the top squark rest frame, and consequently will assume most of the momentum in the lab frame. For the strip with the lowest LSP mass on the other hand, the LSP and the charm quark have about equal mass. After boosting to the lab frame, they will thus share the momentum about equally. As the LSP has a lower momentum in this case, this results in a softer $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum and therefore a lower R^2 value, the final result being a substantially reduced efficiency. To illustrate this, we show the $E_{\mathrm{T}}^{\mathrm{miss}}$ and R^2 distributions for two example signal points in Fig. 7.58.

The efficiencies for mass points of the T1t1t model are lower than for the equivalent points of the T1ttcc models. There are two main reasons for this. Firstly, the top quark is heavier than the charm quark. The afore-mentioned effect will thus occur not only for the lowest LSP mass, but for all masses smaller than the top quark mass. In addition, the combination of the large top quark mass and the considered $\Delta m = 175\,\mathrm{GeV}$ results in a much smaller mass gap between gluino and top squark compared to the T1ttcc models. This in turn results in less boosted top quarks, as can be seen in Fig. 7.4, and thus fewer events passing the boosted W boson tagger.

As no discrepancy from the SM prediction was observed, we compute upper limits on the allowed cross section across the parameter space of the T1ttcc and T1t1t models using the CLs method. Figure 7.61 shows the observed and

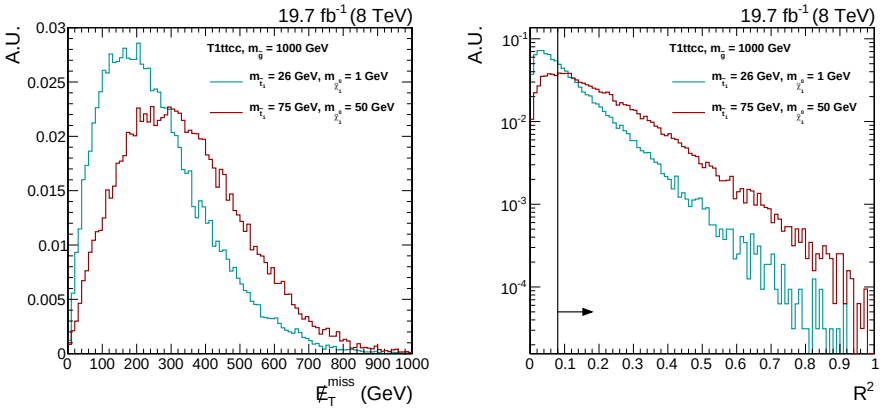

Figure 7.58: Comparison between a T1ttcc signal point with $m_{\tilde{\chi}_1^0} = 1\,\text{GeV}$ and with $m_{\tilde{\chi}_1^0} = 50\,\text{GeV}$. [left] Distribution of $E_{\text{T}}^{\text{miss}}$ with no selection applied. [right] Distribution of R^2. The presence of at least two jets was required in order to calculate the R^2 variable. The arrow indicates the baseline selection requirement of $\text{R}^2 > 0.08$.

expected limit for the T1ttcc model with $\Delta m = 10, 25, 80\,\text{GeV}$ and for the T1t1t model. The uncertainty band for the observed limit is computed by varying the signal cross section within the theoretical uncertainties, of which the PDFs are the main component. The uncertainty on the expected limit incorporates all experimental uncertainties. From these figures we observe that this analysis has made significant inroads into the parameter space of the T1ttcc model. Gluinos with mass up to about 1 TeV have been excluded for neutralinos with mass less than about 500 GeV, when the top squark decays to a charm and a neutralino and $\Delta m < 80\,\text{GeV}$. This also means that top squarks with masses up to about 500 GeV have been excluded for small mass differences with the LSP, given the existence of a gluino with mass less than about 1 TeV. Similarly, for the T1t1t model, top squarks with a mass up to about 450 GeV have been excluded for the scenarios with $\Delta m = 175\,\text{GeV}$ and gluino mass less than 850 GeV. Considering again our cartoon from the introduction in Section 7.1, we have now filled in several of the gaps, as we had set out to do. This is illustrated in Fig. 7.62.

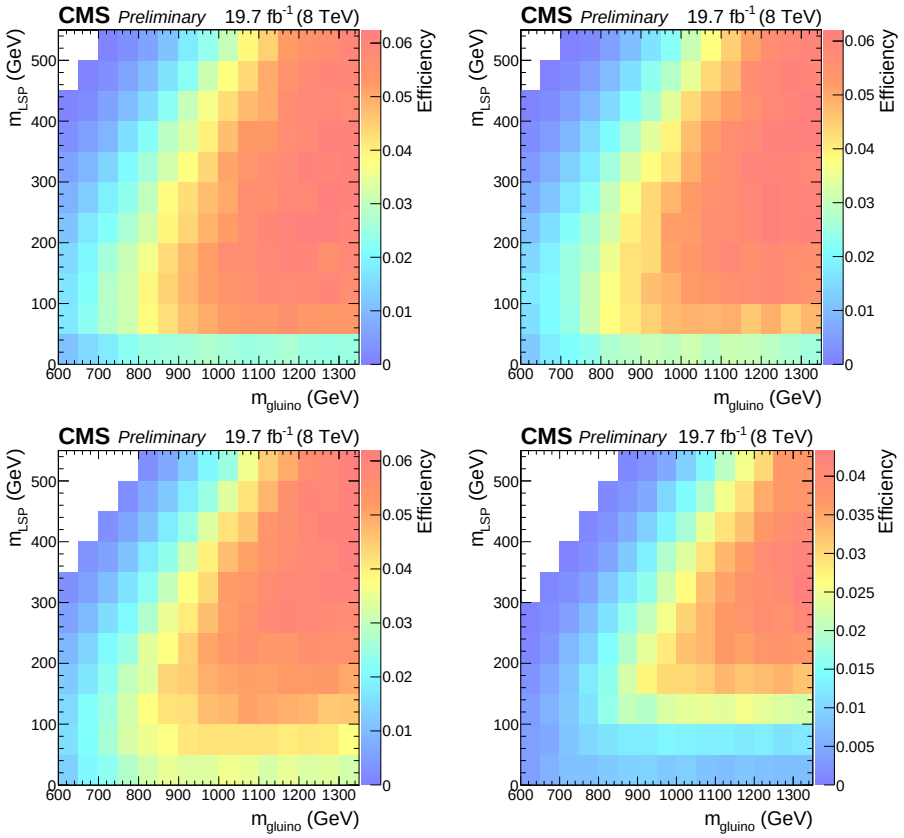

Figure 7.59: Signal region efficiency for the T1ttcc and T1t1t simplified models. Three mass splittings between top squark and LSP are considered for the T1ttcc model: 10, 25 and 80 GeV, shown on the top left, top right, and bottom left, respectively. The T1t1t model is shown on the bottom right plot.

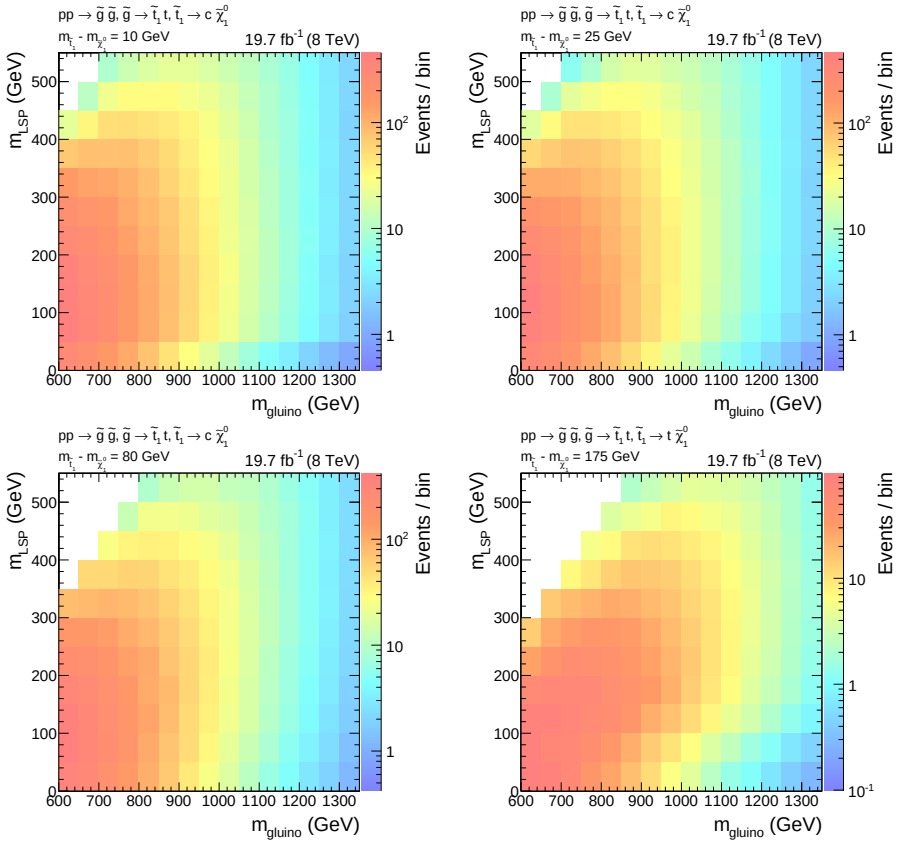

Figure 7.60: Expected signal yield in the signal region for the T1ttcc and T1t1t simplified models. Three mass splittings between top squark and LSP are considered for the T1ttcc model: 10, 25 and 80 GeV, shown on the top left, top right, and bottom left, respectively. The T1t1t model is shown on the bottom right plot.

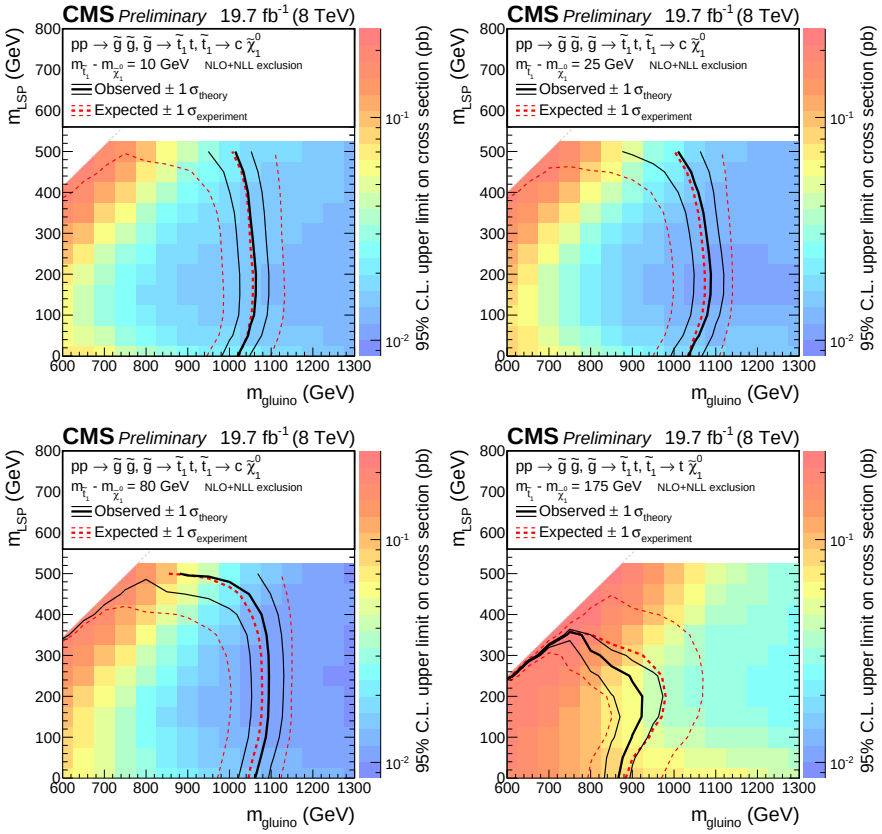

Figure 7.61: Observed and expected limit using CLs for the T1ttcc $\Delta m = 10, 25, 80$ GeV and T1t1t models (top left, top right, bottom left and bottom right, respectively).

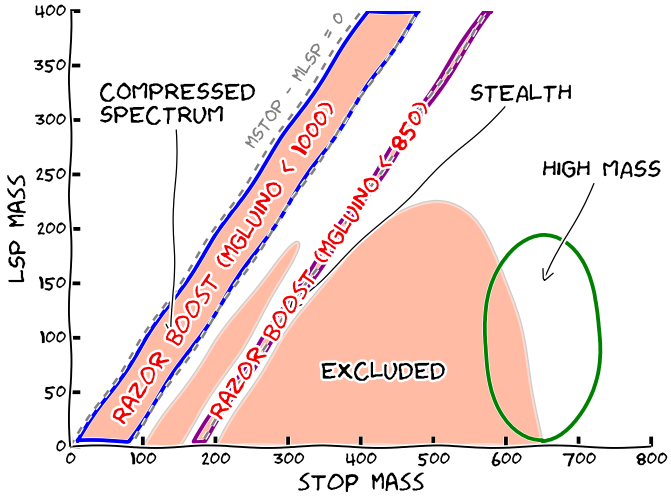

Figure 7.62: Impact of the razor boost analysis on the exclusion limits for top squarks. Assuming the existence of gluinos, the gaps have been closed for certain ranges of gluino masses.

Chapter 8

Summary and discussion

The Standard Model of particle physics encapsulates our current knowledge of the elementary particles and their interactions. It was developed as a quantum field theory over the past fifty years and has been tested thoroughly by many different collider and non-collider experiments. So far, no significant deviations have been observed, and until recently the Higgs boson was the only particle predicted to exist that had not been found. When that elusive particle was finally discovered in 2012, this was the ultimate victory for the Standard Model, and the mechanism of spontaneous electroweak symmetry breaking.

Regardless of all its successes, the Standard Model cannot be the final theory describing the workings of nature. There are several reasons for this, the principal one being that gravity is as yet not included. The Standard Model also cannot explain dark matter, or dark energy. With the discovery of the Higgs boson, the hierarchy problem has become a pertinent issue.

At the core of the hierarchy problem is the question why the electroweak scale is so much smaller than the Planck scale. Radiative corrections to the Higgs boson mass are quadratically dependent on the ultraviolet cutoff scale that is used to regulate the loop diagrams. If the Standard Model is valid up to the point where we need a theory of quantum gravity, then this cutoff is the Planck scale. This then results in a huge correction to the Higgs boson mass. Within the Standard Model this is technically not an issue, but it would require an extremely large amount of finetuning between the bare Higgs mass and the correction to result in the observed Higgs boson mass of 125 GeV. This amount of finetuning is generally viewed to be unnatural, and therefore unwanted.

Several models beyond the Standard Model have been developed in an attempt to address these and other issues. The most popular extension of the Standard Model nowadays is supersymmetry. Supersymmetry adds another symmetry to the theory, a symmetry between fermions and bosons, which results in the introduction of a superpartner for every Standard Model particle. Many times before, symmetries have guided physicists to a deeper understanding of nature. The elegance of supersymmetry is thus definitely part of the appeal, but it is not the only nice feature.

Many supersymmetric models assume the conservation of R-parity, which

leads to a lightest supersymmetric particle (LSP) that is stable, and is a possible dark matter candidate. Supersymmetry, and in particular 'natural supersymmetry', can provide a much more natural solution to the hierarchy problem, removing most of the finetuning. For this to happen the supersymmetric partners of the top quark and the gluon, the top squark and the gluino, must be relatively light, less than about 1 TeV and 1.5 TeV, respectively. These particles, if they exist, should thus be accessible at the collision energies of the LHC.

The possibility that the top squark could be light has motivated several dedicated searches by the CMS and ATLAS collaborations for the direct production of top squark pairs. The sensitivity of many of these searches diminishes when the mass of the top squark approaches that of the LSP, or when the mass difference between the top squark and the LSP is comparable to the top quark mass. The razor boost analysis presented in this thesis aims to fill these gaps in the sensitivity by using the naturalness argument and assuming the existence of an accessible gluino.

In the razor boost analysis, we consider gluino pair production in which the gluino decays to a top squark and a top quark. In the models considered, the gluino is assumed to have a mass around 1-1.5 TeV and the lighter top squark has a mass of a few hundred GeV. The considered top squark decay depends on the assumed mass difference between top squark and LSP. For small mass differences, the top squark decays to a charm quark and the LSP, while for mass differences around the top quark mass the top squark decays to a top quark and the LSP. The models used are simplified models, meaning that all other possible supersymmetric particles have a very large mass and are thus entirely decoupled from the gluino, top squark and LSP. The razor boost analysis described in this thesis is the first analysis within CMS to explicitly probe the gluino-mediated production of top squarks decaying to a charm quark and the LSP, and will thus provide new information on natural supersymmetry.

The razor boost analysis uses the razor kinematic variables M_R and R^2 as main search tool. These variables can discriminate processes with new heavy particles and missing transverse momentum from Standard Model processes. The characteristic mass scale of the new physics is estimated in two ways, using longitudinal information (M_R) and transverse information (R^2) in an event. Signal events are expected to appear as a peak near the characteristic mass scale over an exponentially falling background. The search for new physics is performed in 25 search bins across the high M_R - high R^2 region.

Owing to the significant mass gap presumed to exist between the gluino and the top squark, the top quark from the gluino to top squark decay will receive a large boost. When the boost is large enough, around 700 GeV, the top quark decay products will merge. Since this boost is hard to reach with proton-proton collisions at 8 TeV, we opted to consider boosted W bosons instead, for which merged decay products arise when the boost is around 320 GeV. The hadronically decaying boosted W bosons are identified using jet substructure techniques, in particular jet pruning and N-subjettiness. Pruned jets allow the proper reconstruction of the jet mass, which should be around the W boson mass for merged jets from a W boson. N-subjettiness is a very useful

observable to check the compatibility of a jet having N subjets. For the razor boost analysis, which is the first analysis in the CMS SUSY group to use jet substructure, we require that the boosted W boson candidates should be consistent with having two subjets.

The signal region selection requires the presence of at least one boosted W boson candidate and one jet tagged as originating from a b quark. We also require the presence of three jets, and veto events with leptons. The background from Standard Model processes in the 25 signal region bins is estimated using observations in data control regions and scale factors, calculated from simulated data, that relate the number of events in one region to that in another. Three control regions, Q, W and T, are defined to select high-purity samples of multijet, $W(\to \ell\nu)$+jets and $t\bar{t}$ processes, respectively. The scale factors are global scale factors, meaning that they are integrated over the full (M_R,R^2) space. The reason for this is insufficient simulated events at high M_R and high R^2.

The background estimation method uses a likelihood-based approach. The likelihood for each search bin is modelled as a Poisson. The expected background components in the signal bins, corresponding to the $t\bar{t}$, $W(\to \ell\nu)$+jets and multijet processes, are constrained using a prior distribution, which is the translation of the relationships between signal and control regions, and incorporates all statistical and systematic uncertainties. The systematic uncertainties are sampled simultaneously, which ensures that any correlations are taken into account automatically. The razor boost analysis is the first SUSY analysis within CMS that uses this complete treatment of the systematic uncertainties.

The background prediction result for each signal bin is derived from the prior distribution. The results are found to be in agreement with the observation in data. We can thus conclude that there is no evidence of new physics in the phase space that was probed by the razor boost analysis. These results can also be used to constrain the allowed parameter space for more specific new physics scenarios. As mentioned earlier, we consider two simplified models of gluino pair production, with the gluino decaying to a top squark and a top quark. For the case where the top squark and the LSP have a small mass difference ($\leq 80\,\text{GeV}$), the top squark is assumed to decay to a charm quark and the LSP. In this model, we can exclude gluinos up to about 1 TeV for LSP masses up to about 500 GeV. Or equivalently, we can exclude top squarks decaying to charm and LSP with masses up to about 500 GeV, provided that there is a gluino with mass less than about 1 TeV. Similarly, for the model in which the top squark decays to a top quark and the LSP, with a mass difference of 175 GeV between them, we have excluded top squarks with a mass up to about 450 GeV if the gluino mass is less than 850 GeV.

The weakest point of the razor boost analysis is the use of global scale factors to translate between signal and control regions, as there could be shape differences that are not accounted for. A set of closure tests was performed to check whether the global scale factors were adequate for the precision achieved in the analysis. We found that an extra uncertainty of 33% needed to be applied to the multijet prediction in order to cover for possible shape differences. Once applied, there was no further evidence of a shape difference within uncertainties.

The decision to use global scale factors was driven by the lack of statistical precision in the simulated samples. This can thus be improved upon in the future by having more simulated events in the boosted regime.

The razor boost analysis features two techniques that set it apart from other SUSY analyses: the use of boosted W boson tagging, and the statistical treatment of the systematic uncertainties. Both of these techniques will be very valuable for analyses with the 13 TeV dataset to be collected from 2015 onwards.

Because of the increase in centre-of-mass energy from 8 to 13 TeV, we expect the produced particles to be more boosted in general. The use of jet substructure techniques will thus be a powerful addition to many analyses. For the razor boost analysis the raised energy will mean that the signal efficiency will also become larger, boosting the reach of the analysis beyond what would be expected from the increased cross section alone.

For the razor boost analysis at 13 TeV, adding extra channels could help increase the sensitivity to a wider range of models. The first option, which I believe will be very beneficial, is to include boosted top tagging. Having a well-thought-out strategy that combines final states with both boosted W bosons and boosted top quarks will ensure that a wide range of mass spectra can be probed. The boosted top signal region will be optimal for very large mass splittings, whereas the boosted W region will be more targeted towards medium mass splittings. Secondly, the leptonic channels could be added to the analysis. Boosted W bosons that decay leptonically will feature leptons with very large transverse momentum. These leptons can be very close to the b jet from the top quark decay, and could thus fail the isolation requirements. A dedicated treatment of these leptons is thus necessary.

Appendix A

Data sets

Table A.1: Simulated samples for the various background processes. The dataset names have been stripped from /AODSIM.

Process	Dataset name	σ (pb)	Level
$t\bar{t}$	/TTJets_MassiveBinDECAY_TuneZ2star_8TeV-madgraph-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1 /TTJets_MassiveBinDECAY_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V7C_FSIM-v2	245.8	NNLO
QCD multijet, $100 < H_T < 250$ GeV	/QCD_HT-100To250_TuneZ2star_8TeV-madgraph-pythia/Summer12_DR53X-PU_S10_START53_V7A-v1	10.367×10^7	LO
QCD multijet, $250 < H_T < 500$ GeV	/QCD_HT-250To500_TuneZ2star_8TeV-madgraph-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1	276000	LO
QCD multijet, $500 < H_T < 1000$ GeV	/QCD_HT-500To1000_TuneZ2star_8TeV-madgraph-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1	8426	LO
QCD multijet, $H_T > 1000$ GeV	/QCD_HT-1000ToInf_TuneZ2star_8TeV-madgraph-pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1	204	LO
$W(\to \ell\nu)$+jets, $250 < H_T < 300$ GeV	/WJetsToLNu_HT-250To300_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1	48.01	LO
$W(\to \ell\nu)$+jets, $300 < H_T < 400$ GeV	/WJetsToLNu_HT-300To400_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1	38.3	LO
$W(\to \ell\nu)$+jets, $H_T > 400$ GeV	/WJetsToLNu_HT-400Toinf_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1	25.22	LO
t, tW-channel	/T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	11.1	NNLO
$\bar{t}$, tW-channel	/Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	11.1	NNLO
t, s-channel	/T_s-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	3.79	NNLO
$\bar{t}$, s-channel	/Tbar_s-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	1.76	NNLO
t, t-channel	/T_t-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	56.4	NNLO
$\bar{t}$, t-channel	/Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	30.7	NNLO
$Z(\to \nu\bar{\nu})$+jets, $50 < H_T < 100$ GeV	/ZJetsToNuNu_50_HT_100_TuneZ2Star_8TeV_madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	381.2	LO
$Z(\to \nu\bar{\nu})$+jets, $100 < H_T < 200$ GeV	/ZJetsToNuNu_100_HT_200_TuneZ2Star_8TeV_madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	160.3	LO
$Z(\to \nu\bar{\nu})$+jets, $200 < H_T < 400$ GeV	/ZJetsToNuNu_200_HT_400_TuneZ2Star_8TeV_madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	41.49	LO
$Z(\to \nu\bar{\nu})$+jets, $H_T > 400$ GeV	/ZJetsToNuNu_400_HT_inf_TuneZ2Star_8TeV_madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	5.274	LO
$Z/\gamma^*(\to \ell\bar{\ell})$+jets, $200 < H_T < 400$ GeV	/DYJetsToLL_HT-200To400_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	19.73	LO
$Z/\gamma^*(\to \ell\bar{\ell})$+jets, $H_T > 400$ GeV	/DYJetsToLL_HT-400ToInf_TuneZ2Star_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	2.826	LO
$W(\to q\bar{q}) + b\bar{b}$	/Wbb_FullyHadronic_8TeV_madgraph/Summer12_DR53X-PU_S10_START53_V7C-v1	121	LO
$Z/\gamma^* \to c\bar{c}$	/DYToCC_M_50_TuneZ2star_8TeV_pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1	3060.099	LO
$Z/\gamma^* \to b\bar{b}$	/DYToBB_M_50_TuneZ2star_8TeV_pythia6/Summer12_DR53X-PU_S10_START53_V7A-v1	3840.386	LO

Table A.2: Simulated samples for the various background processes. The dataset names have been stripped from /AODSIM.

Process	Dataset name	σ (pb)	Level
WW	/WW_TuneZ2star_8TeV_pythia6_tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	54.838	LO
WZ	/WZ_TuneZ2star_8TeV_pythia6_tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	32.3161	LO
ZZ	/ZZ_TuneZ2star_8TeV_pythia6_tauola/Summer12_DR53X-PU_S10_START53_V7A-v1	8.258	LO
WWW	/WWW_8TeV-aMCatNLO-herwig/Summer12_DR53X-PU_S10_START53_V7C-v1	0.08058	NLO
WWZ	/WWZNoGstarJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	0.05795	NLO
WZZ	/WZZ_8TeV-aMCatNLO-herwig/Summer12_DR53X-PU_S10_START53_V7C-v1	0.01968	NLO
ZZZ	/ZZZNoGstarJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	0.005527	NLO
$WW\gamma$	/WWGJets_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1	0.528	LO
$t\bar{t}Z$	/ttbarZ_8TeV-Madspin_aMCatNLO-herwig/Summer12_DR53X-PU_S10_START53_V19-v1	0.2057	NLO
$t\bar{t}W$	/TTbarW_8TeV-aMCatNLO-herwig/Summer12_DR53X-PU_S10_START53_V7C-v1	0.232	NLO
$t\bar{t}\gamma$+jets	/TTGJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V19-v1	1.44	LO
$t\bar{t}WW$+jets	/TTWWJets_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7A-v1	0.002037	LO

Table A.3: MC simulated samples for the signal events.

Topology	Dataset name
T2tt $pp \to \tilde{t}_1\tilde{t}_1^*$, $\tilde{t}_1 \to t\tilde{\chi}_1^0$	/SMS-T2tt_mStop-150to350_mLSP-0to250_8TeV-Pythia6Z/Summer12-START52_V9_FSIM-v1 /SMS-T2tt_mStop-375to475_mLSP-0to375_8TeV-Pythia6Z/Summer12-START52_V9_FSIM-v1 /SMS-T2tt_mStop-500to650_mLSP-0to225_8TeV-Pythia6Z/Summer12-START52_V9_FSIM-v1 /SMS-T2tt_mStop-500to650_mLSP-250to550_8TeV-Pythia6Z/Summer12-START52_V9_FSIM-v1 /SMS-T2tt_mStop-675to800_mLSP-0to275_8TeV-Pythia6Z/Summer12-START52_V9_FSIM-v1 /SMS-T2tt_mStop-675to800_mLSP-300to700_8TeV-Pythia6Z/Summer12-START52_V9_FSIM-v1 /SMS-8TeV-Pythia6Z_T2tt_mStop-150to475_mLSP-1/Summer12-START52_V9_FSIM-v1 /SMS-8TeV-Pythia6Z_T2tt_mStop-500to800_mLSP-1/Summer12-START52_V9_FSIM-v1 /SMS-T2tt_mStop-825to900_mLSP-1_and_mLSP-25to800_8TeV-Pythia6Zstar/Summer12-START52_V9_FSIM-v3 /SMS-T2tt_mStop-925to1000_mLSP-1_and_mLSP-25to900_8TeV-Pythia6Zstar/Summer12-START52_V9_FSIM-v3
T1ttcc, fixed gluino mass $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to t\tilde{t}_1$, $\tilde{t}_1 \to c\tilde{\chi}_1^0$	/SMS-MadGraph_Pythia6Zstar_8TeV_T1ttcc_2J_mGo-1000_mStop-310-325-350-375_mLSP-300/Summer12-START52_V9_FSIM-v1 /SMS-MadGraph_Pythia6Zstar_8TeV_T1ttcc_2J_mGo-1000_mStop-410-425-450-475_mLSP-400/Summer12-START52_V9_FSIM-v1 /SMS-MadGraph_Pythia6Zstar_8TeV_T1ttcc_2J_mGo-1000_mStop-510-525-550-575_mLSP-500/Summer12-START52_V9_FSIM-v1 /SMS-MadGraph_Pythia6Zstar_8TeV_T1ttcc_2J_mGo-1000_mStop-610-625-650-675_mLSP-600/Summer12-START52_V9_FSIM-v1 /SMS-MadGraph_Pythia6Zstar_8TeV_T1ttcc_2J_mGo-1000_mStop-710-725-750-775_mLSP-700/Summer12-START52_V9_FSIM-v1 /SMS-MadGraph_Pythia6Zstar_8TeV_T1ttcc_2J_mGo-1000_mStop-810-825-850-875_mLSP-800/Summer12-START52_V9_FSIM-v1
T1ttcc, fixed $m_{\tilde{t}_1} - m_{\tilde{\chi}_1^0}$ $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to t\tilde{t}_1$, $\tilde{t}_1 \to c\tilde{\chi}_1^0$	/SMS-T1ttcc_2J_mGo-600to750_mLSP-1to500_DM-10_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V19_FSIM_PU_S12-v1 /SMS-T1ttcc_2J_mGo-600to750_mLSP-1to500_DM-25_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V19_FSIM_PU_S12-v1 /SMS-T1ttcc_2J_mGo-600to750_mLSP-1to500_DM-80_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V19_FSIM_PU_S12-v1 /SMS-T1ttcc_2J_mGo-800to1300_mLSP-1to500_DM-10_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V19_FSIM_PU_S12-v1 /SMS-T1ttcc_2J_mGo-800to1300_mLSP-1to500_DM-25_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V19_FSIM_PU_S12-v1 /SMS-T1ttcc_2J_mGo-800to1300_mLSP-1to500_DM-80_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V19_FSIM_PU_S12-v1
T1t1t $pp \to \tilde{g}\tilde{g}$, $\tilde{g} \to t\tilde{t}_1$, $\tilde{t}_1 \to t\tilde{\chi}_1^0$	/SMS-T1t1t_2J_mGo-800to1300_mLSP-1to500_DM-175_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V19_FSIM_PU_S12-v1 /SMS-T1t1t_2J_mGo-600to750_mLSP-1to500_DM-175_TuneZ2star_8TeV-madgraph-tauola/Summer12-START53_V19_FSIM_PU_S12-v1

Appendix B

Systematic uncertainties across the T1ttcc ($\Delta m = 25\,\mathrm{GeV}$) mass plane

In the section on systematic uncertainties, in Table 7.14, we presented the average of the systematic effects across the signal mass plane. Here, in Figs. B.1 to B.9, we show the one standard deviation (σ) up and down variations of the considered signal systematics for each point in the ($m_{\tilde{g}}$,$m_{\tilde{\chi}_1^0}$) plane.

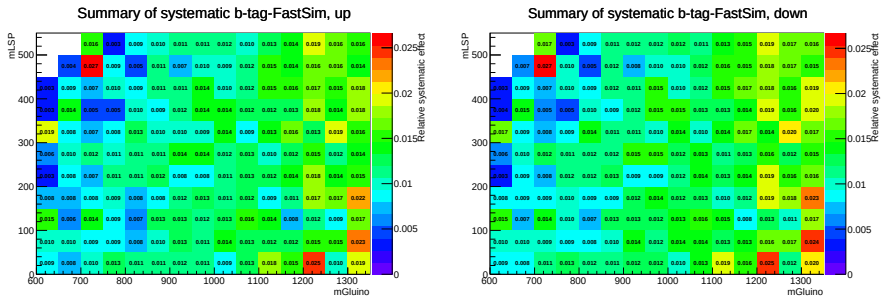

Figure B.1: 1σ up (left) and down (right) variation for b tag FullSim/Fastsim SF.

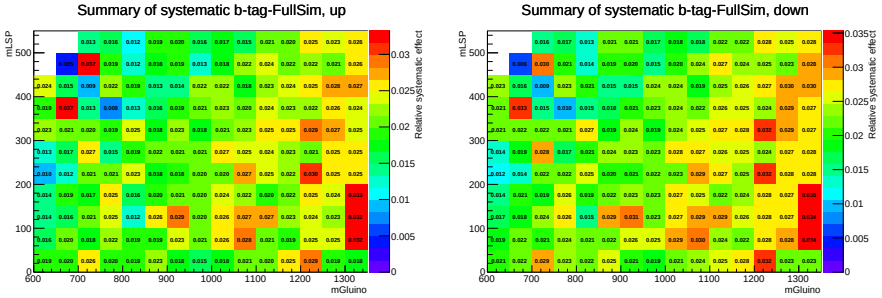

Figure B.2: 1σ up (left) and down (right) variation for b tag Data/FullSim SF.

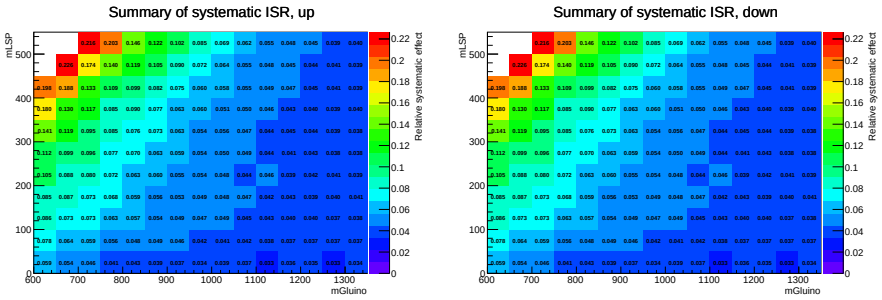

Figure B.3: 1σ up (left) and down (right) variation for ISR reweighting.

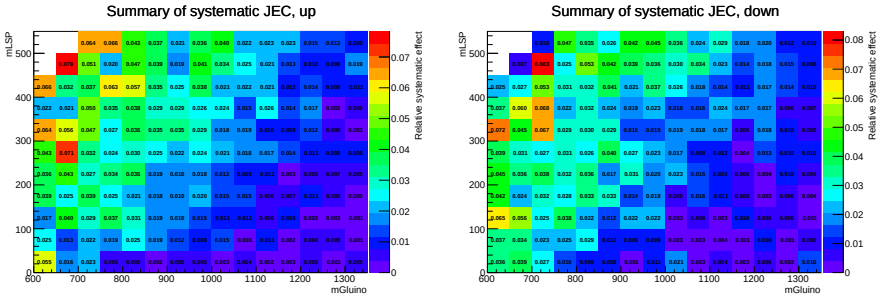

Figure B.4: 1σ up (left) and down (right) variation for jet energy corrections.

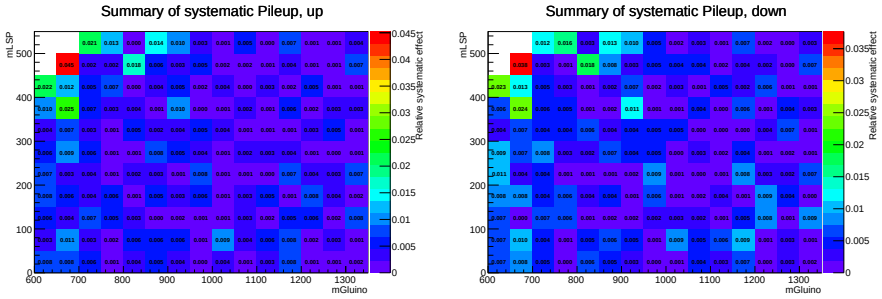

Figure B.5: 1σ up (left) and down (right) variation for pileup reweighting.

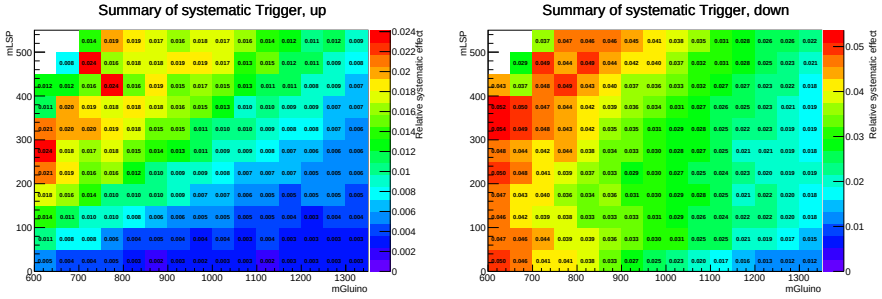

Figure B.6: 1σ up (left) and down (right) variation for trigger efficiency.

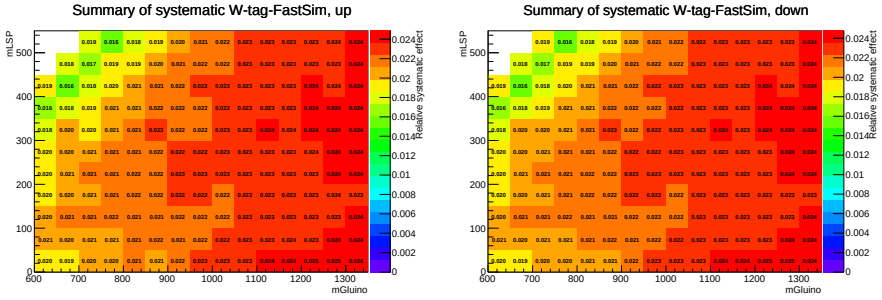

Figure B.7: 1σ up (left) and down (right) variation for W boson tag FullSim/FastSim SF.

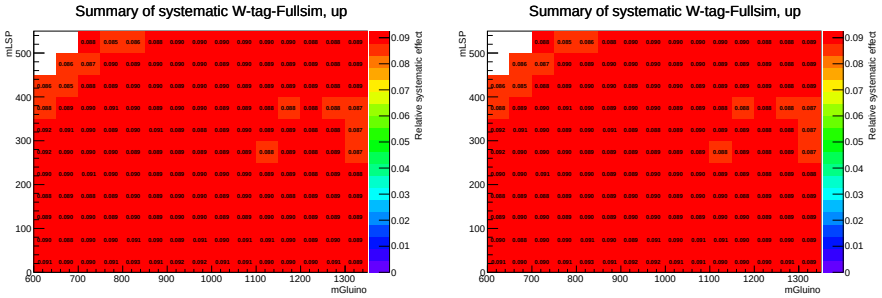

Figure B.8: 1σ up (left) and down (right) variation for W boson tag Data/FullSim SF.

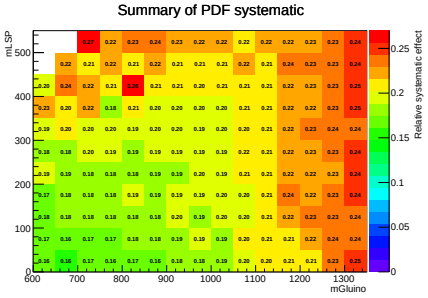

Figure B.9: 1σ variation for the parton distribution functions.

Nederlandstalige samenvatting

Het Standaard Model van de deeltjesfysica beschrijft onze huidige kennis van de elementaire deeltjes en hun interacties. Het werd ontwikkeld als een kwantumveldentheorie gedurende de laatste vijftig jaar, en is grondig getest door een groot aantal experimenten, al dan niet aan een deeltjesversneller. Tot nu toe zijn er nog geen significante afwijkingen geobserveerd, en tot voor kort was het Higgs boson het enige voorspelde deeltje dat nog niet gevonden was. Toen dat moeilijk te vangen deeltje eindelijk ontdekt werd in 2012, was dit de ultieme overwinning voor het Standaard Model en het mechanisme van de spontane elektrozwakke symmetriebreking.

Ongeacht het grote succes van het Standaard Model, kan dit toch niet de definitieve theorie van de natuur zijn. Hiervoor zijn verschillende redenen, in het bijzonder kan het Standaard Model de zwaartekracht niet beschrijven. Het model bevat ook geen deeltjes die als kandidaat voor donkere materie kunnen dienen, en er is ook geen verklaring voor donkere energie. Nu het Higgsdeeltje ontdekt is, is het zogenaamde hiërarchieprobleem ook heel relevant geworden.

De kern van het hiërarchieprobleem ligt bij de vraag waarom de electrozwakke schaal zo veel kleiner is dan de Planckschaal. Kwantumcorrecties tot de massa van het Higgsdeeltje zijn namelijk kwadratisch afhankelijk van de cutoff schaal die gebruikt wordt om de loop-diagrammen te regularizeren. Als het Standaard Model geldig is tot op het punt waar een theorie van kwantumgravitatie nodig is, dan is deze cutoff schaal gelijk aan de Planckschaal. Dit zal dan op zijn beurt resulteren in een enorme correctie van de Higgs boson massa. Technisch gezien is dit geen probleem, maar het betekent dat er een zeer grote graad van finetuning aanwezig moet zijn tussen de naakte massa en de correctie om tot de geobserveerde waarde van 125 GeV te leiden. Deze graad van finetuning wordt algemeen aanzien als onnatuurlijk, en dus ongewenst.

Verschillende modellen voor nieuwe fysica zijn ontwikkeld in een poging om deze en andere gebreken op te lossen. De dag van vandaag is supersymmetrie de meest populaire theorie. Supersymmetrie voegt een nieuwe fundamentele symmetrie toe, één tussen fermionen en bosonen. Het resultaat hiervan is de invoering van een superpartnerdeeltje voor elk bestaand deeltje in het Standaard Model. Doorheen de geschiedenis van de deeltjesfysica hebben symmetrieën ons telkens opnieuw geleid tot een dieper begrip van de natuur. De elegantie van supersymmetrie is dus zeker een aantrekkingspunt, maar dit is niet het

enige.

Veel supersymmetrische modellen nemen aan dat de R-pariteit behouden is. Hieruit volgt dat het lichtste supersymmetrische deeltje (LSP) stabiel is, en bovendien een kandidaat is voor donkere materie. Supersymmetrie, en in het bijzonder de zogenaamde 'natuurlijke supersymmetrie', kan ook een meer natuurlijke verklaring voor het hiërachieprobleem bieden, waarbij de finetuning grotendeels verwijderd wordt. Opdat dit zou gebeuren, moeten de superpartners van de top quark en het gluon, de top squark en het gluino, relatief licht zijn, respectievelijk lichter dan 1 TeV en 1.5 TeV. Als deze deeltjes bestaan, zouden ze dus gevonden kunnen worden in de botsingen van de LHC.

De mogelijkheid dat de top squark licht kan zijn, diende als motivatie voor verschillende analyses in de CMS en ATLAS experimenten, waarbij er specifiek gezocht wordt naar de directe productie van top squark paren. De gevoeligheid van deze analyses verkleint wanneer de massa van de top squark dat van het LSP benadert, of wanneer het verschil in hun massa's ongeveer de top quark massa bedraagt. De razor boost analyse, die in deze thesis voorgesteld wordt, heeft als doel deze gaten in de gevoeligheid op te vullen. In lijn met de voorspellingen van de natuurlijke supersymmetrie wordt het bestaan van een relatief licht gluino verondersteld.

In de razor boost analyse beschouwen we de paar-productie van gluinos, die vervolgens vervallen naar een top squark en een top quark. De gluinos hebben een massa van ongeveer 1-1.5 TeV, en de top squarks hebben een massa van enkele honderden GeV. Het verval van de top squark hangt af van het verondersteld massaverschil tussen de top squark en het LSP. Voor kleine massaverschillen vervalt de top squark naar een charm quark en het LSP, terwijl voor massaverschillen rond de top quark massa de top squark naar een top quark en het LSP vervalt. De gebruikte modellen zijn zogenaamde 'simplified models', wat zoveel betekent als dat alle andere superpartners een heel grote massa hebben en dus niet bijdragen tot de productie en het verval van de gluinos, top squarks en LSP's. De razor boost analyse is de eerste analyse binnen CMS die expliciet de productie van top squarks via gluinos, en vervallend naar een charm quark en het LSP, onderzoekt. Deze analyse zal dus nieuwe information verschaffen over supersymmetrie.

De razor kinematische variabelen, M_R en R^2, zijn de belangrijkste variabelen voor de razor boost analyse. Deze variabelen laten toe om een onderscheid te maken tussen de gewone processen uit het Standaard Model, en de processen waarbij nieuwe, zware deeltjes geproduceerd worden die vergezeld gaan van een hoeveelheid ontbrekende energie. De karakteristieke massaschaal van de nieuwe deeltjes wordt geschat op twee manieren, door gebruik te maken van longitudinale informatie enerzijds (M_R), en transversale informatie anderzijds (R^2). Het gezochte signaal zal verschijnen als een piek bovenop een exponentieel vallende achtergrond. De zoektocht naar nieuwe fysica die hier is beschreven, wordt uitgevoerd in 25 gebieden in het tweedimensionale (M_R, R^2) vlak.

Dankzij het verondersteld grote massaverschil tussen het gluino en de top squark, zal de top quark uit het verval van de gluino een grote boost krijgen. Als deze boost groot genoeg wordt, ongeveer 700 GeV, dan zullen de vervalproducten van de top quark gecollimeerd worden, en met elkaar overlap-

pen in de detector. Aangezien deze boost moeilijk te bereiken is met proton-proton botsingen van 8 TeV, zullen we in plaats daarvan gebooste W bosonen beschouwen. Door hun lagere massa overlappen de vervalproducten van W bosonen al met elkaar bij een boost van ongeveer 320 GeV. W bosonen die hadronisch vervallen worden geïdentificeerd door gebruik te maken van technieken die de substructuur van jets bestuderen, in het bijzonder 'jet pruning' en 'N-subjettiness'. De eerste van die technieken laat toe om de massa van jets nauwkeurig te bepalen. De gezochte jets moeten dus een massa hebben die overeenstemt met de bekende massa van een W boson. De N-subjettiness observabelen testen of een jet compatibel is met het hebben van N subjets. Voor de razor boost analyse zullen we dus vereisen dat de W boson kandidaten consistent zijn met de hypothese van twee subjets.

De selectie van het mogelijke signaal omvat de vereiste dat er minstens één W boson kandidaat is, en minstens één jet die van een b quark afkomstig is. We verwijderen alle botsingen waarbij leptonen geproduceerd werden. De achtergrond van gekende processen uit het Standaard Model wordt voor elk van de 25 regio's berekend via observaties in controlegebieden en schaalfactoren die de relatie tussen het signaalgebied en de controlegebieden beschrijven. Er werden drie controlegebieden gedefinieerd, Q, W en T, die respectievelijk de multijet, $t\bar{t}$ en $W(\to \ell\nu)$+jets processen selecteren. Als gevolg van het tekort aan gesimuleerde data voor grote waarden van M_R en R^2, gebruiken we globale schaalfactoren, ze zijn dus geïntegreerd over het volledige (M_R, R^2) vlak.

De methode om de achtergrond te berekenen is gebaseerd op een likelihood. De likelihood voor elke signaalregio is een Poisson. De verwachtte achtergrond-componenten in de signaalgebieden, overeenstemmend met de multijet, $t\bar{t}$ en $W(\to \ell\nu)$+jets processen, worden gemodelleerd door een 'prior distribution'. Deze distributie omvat de relaties tussen de signaal- en controlegebieden, en all statistische en systematische onzekerheden. Alle systematische onzekerheden worden gelijktijdig gevarieerd, wat tot gevolg heeft dat alle mogelijke correlaties automatisch in rekening gebracht worden. De razor boost analyse is de eerste analyse binnen de supersymmetrie-groep van CMS die deze aanpak gebruikt heeft.

Het resultaat van de voorspelling van het aantal achtergrondgebeurtenissen in elke signaalregio werd afgeleid van de prior, en is overeenstemming met de observatie in de data. We kunnen dus besluiten dat er geen aanwijzingen zijn voor het bestaan van nieuwe fysica in de faseruimte die onderzocht werd door de razor boost analyse. We kunnen deze resultaten ook gebruiken om limieten te zetten op parameters van specifieke modellen. Zoals al eerder vermeld, beschouwen we twee modellen van gluino paarproductie, waarbij de gluinos vervallen naar een top squark en een top quark. Voor het model waarbij de top squark op zijn beurt vervalt naar een charm quark en het LSP, en het massaverschil tussen de top squark en het LSP klein is (≤ 80 GeV), kunnen we de aanwezigheid van gluinos met een massa tot 1 TeV uitsluiten, zolang het LSP een massa heeft die kleiner is dan 500 GeV. We kunnen dit ook omdraaien en besluiten dat we top squarks met een massa tot ongeveer 500 GeV kunnen uitsluiten, op voorwaarde dat ze vervallen tot een charm quark en het LSP, en dat er een gluino aanwezig is in het spectrum met een massa van maximaal 1 TeV.

Voor het andere model kunnen we een gelijkaardige vaststelling maken. Top squarks die vervallen naar een top quark en het LSP, waarbij het massaverschil gelijk is aan de top quark massa, kunnen uitgesloten worden als hun massa lager is dan ongeveer 450 GeV, en er gluinos bestaan met een maximale massa van 850 GeV.

Het zwakste punt van de razor boost analyse is het gebruik van globale schaalfactoren om de relaties tussen signaalregio en controleregio's te beschrijven. Er kan namelijk een verschil zijn in de vorm van de distributies die daardoor niet in rekening gebracht wordt. We hebben verschillende tests uitgevoerd om te verifiëren dat de globale schaalfactoren voldoende geschikt waren volgens de statistische precisie van de analyse. Als resultaat van deze testen werd er besloten om een bijkomende onzekerheid van 33% toe te passen op de multijet voorspelling. De beslissing om globale schaalfactoren te gebruiken, werd genomen omdat de statistische precisie van de simulatie niet voldoende was om een aparte schaalfactor voor elke regio te definiëren. In de toekomst kan dit onderdeel van de analyse dus verbeterd worden door meer gesimuleerde data te voorzien in de regio's met hoge boost.

De razor boost analyse gebruikte twee technieken die de analyse onderscheid van andere supersymmetrie analyses: het identificeren van W bosonen met hoge boost, en de statistische behandeling van de systematische onzekerheden. Deze beide technieken zullen erg waardevol zijn voor toekomstige analyses die de dataset van 2015 zullen gebruiken.

Door het verhogen van de botsingsenergie van 8 TeV naar 13 TeV zullen geproduceerde deeltjes meer geboost worden. Het gebruik van jet substructuur kan dus een groot voordeel opleveren voor veel analyses. Voor de razor boost analyse betekent de verhoogde energie dat de efficiëntie van de selectie ook zal verhogen, meer dan wat verwacht zou worden van de verhoogde werkzame doorsnede alleen.

Het toevoegen van extra signaalgebieden kan ook helpen om de gevoeligheid aan meer verscheidene modellen te verbeteren. Een eerste mogelijkheid is het invoegen van boosted top quark identificatie. Door topologieën met boosted top quarks en boosted W bosonen te combineren, kunnen meer modellen getest worden. Enerzijds zal het signaalgebied met boosted top quarks optimaal zijn voor modellen met grote verschillen in massa tussen de verschillende deeltjes. Anderzijds is het gebruik van boosted W bosons gepast voor modellen met meer gemiddelde massaverschillen. De tweede optie is het toevoegen van een signaalregio waarbij de aanwezigheid van een lepton vereist wordt. W bosonen met een grote boost zullen vervallen naar leptonen met veel transversale impuls. Deze leptonen kunnen zich bovendien dicht bij de b jet van het top quark verval bevinden, wat ervoor kan zorgen dat ze aan bepaalde isolatiecriteria niet voldoen. Een specifieke behandeling van deze leptonen zal dus nodig zijn.

Bibliography

[1] B. Povh, K. Rith, C. Scholz, and F. Zersche, "Particles and nuclei: An Introduction to the physical concepts,".

[2] A. Bettini, *Introduction to elementary particle physics*. Cambridge University Press, 2014.

[3] M. E. Peskin and D. V. Schroeder, "An Introduction to quantum field theory,".

[4] C. Burgess and G. Moore, *The standard model: A primer*. Cambridge University Press, 2007. ISBN-9780521860369.

[5] **Particle Data Group** Collaboration, K. Olive *et al.*, "Review of Particle Physics," *Chin.Phys.* **C38** (2014) 090001.

[6] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys.Rev.Lett.* **13** (1964) 321–323.

[7] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys.Rev.Lett.* **13** (1964) 508–509.

[8] Flip Tanedo. `http://www.quantumdiaries.org/2011/11/21/why-do-we-expect-a-higgs-boson-part-i-electroweak-symmetry-breaking`.

[9] **CMS** Collaboration, S. Chatrchyan *et al.*, "Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV," *Phys.Lett.* **B710** (2012) 26–48, `arXiv:1202.1488 [hep-ex]`.

[10] **CMS** Collaboration, S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.* **B716** (2012) 30–61, `arXiv:1207.7235 [hep-ex]`.

[11] **CMS** Collaboration, V. Khachatryan *et al.*, "Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV," `arXiv:1412.8662 [hep-ex]`. Submitted to Eur. Phys. J. C.

[12] `http://ckmfitter.in2p3.fr/`.

[13] `http://project-gfitter.web.cern.ch/project-gfitter/`.

[14] http://zfitter.com/.

[15] http://lepewwg.web.cern.ch/LEPEWWG/.

[16] **Gfitter Group** Collaboration, M. Baak *et al.*, "The global electroweak fit at NNLO and prospects for the LHC and ILC," *Eur.Phys.J.* **C74** (2014) 3046, arXiv:1407.3792 [hep-ph].

[17] H. Flacher, M. Goebel, J. Haller, A. Hocker, K. Monig, *et al.*, "Revisiting the Global Electroweak Fit of the Standard Model and Beyond with Gfitter," *Eur.Phys.J.* **C60** (2009) 543–583, arXiv:0811.0009 [hep-ph].

[18] P. Van Nieuwenhuizen, "Supergravity," *Phys.Rept.* **68** (1981) 189–398.

[19] D. Z. Freedman and A. Van Proeyen, *Supergravity*. Cambridge University Press, 2012.

[20] H. Nastase, "Introduction to Supergravity," arXiv:1112.3502 [hep-th].

[21] **Planck** Collaboration, P. Ade *et al.*, "Planck 2015 results. XIII. Cosmological parameters," arXiv:1502.01589 [astro-ph.CO].

[22] A. D. Sakharov, "Violation of cp invariance, c asymmetry, and baryon asymmetry of the universe," *JETP Lett.* **5** no. 1, (1967) 24–27.

[23] **KamLAND** Collaboration, S. Abe *et al.*, "Precision Measurement of Neutrino Oscillation Parameters with KamLAND," *Phys.Rev.Lett.* **100** (2008) 221803, arXiv:0801.4589 [hep-ex].

[24] **T2K** Collaboration, K. Abe *et al.*, "Precise Measurement of the Neutrino Mixing Parameter θ_{23} from Muon Neutrino Disappearance in an Off-Axis Beam," *Phys.Rev.Lett.* **112** no. 18, (2014) 181801, arXiv:1403.1532 [hep-ex].

[25] **OPERA** Collaboration, N. Agafonova *et al.*, "Observation of tau neutrino appearance in the CNGS beam with the OPERA experiment," *PTEP* **2014** no. 10, (2014) 101C01, arXiv:1407.3513 [hep-ex].

[26] F. Klinkhamer, "Neutrino mass and the Standard Model," *Mod.Phys.Lett.* **A28** (2013) 1350010, arXiv:1112.2669 [hep-ph].

[27] **ALEPH** Collaboration, D. Decamp *et al.*, "A Precise Determination of the Number of Families With Light Neutrinos and of the Z Boson Partial Widths," *Phys.Lett.* **B235** (1990) 399.

[28] A. Djouadi, "The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model," *Phys.Rept.* **459** (2008) 1–241, arXiv:hep-ph/0503173 [hep-ph].

[29] H.-C. Cheng, "Little Higgs, Non-standard Higgs, No Higgs and All That," arXiv:0710.3407 [hep-ph].

[30] J. Reuter and M. Tonini, "Can the 125 GeV Higgs be the Little Higgs?," *JHEP* **1302** (2013) 077, arXiv:1212.5930 [hep-ph].

[31] M. Schmaltz and D. Tucker-Smith, "Little Higgs review," *Ann.Rev.Nucl.Part.Sci.* **55** (2005) 229–270, arXiv:hep-ph/0502182 [hep-ph].

[32] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, "The Hierarchy problem and new dimensions at a millimeter," *Phys.Lett.* **B429** (1998) 263–272, arXiv:hep-ph/9803315 [hep-ph].

[33] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for signatures of extra dimensions in the diphoton mass spectrum at the Large Hadron Collider," *Phys.Rev.Lett.* **108** (2012) 111801, arXiv:1112.0688 [hep-ex].

[34] J. Wess and B. Zumino, "Supergauge Transformations in Four-Dimensions," *Nucl. Phys. B* **70** (1974) 39.

[35] Y. A. Golfand and E. P. Likhtman, "Extension of the Algebra of Poincare Group Generators and Violation of p Invariance," *JETP Lett.* **13** (1971) 323.

[36] A. H. Chamseddine, R. L. Arnowitt, and P. Nath, "Locally Supersymmetric Grand Unification," *Phys. Rev. Lett.* **49** (1982) 970–982.

[37] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells, "Study of constrained minimal supersymmetry," *Phys. Rev.* **D49** (1994) 6173–6210, arXiv:hep-ph/9312272 [hep-ph].

[38] P. Fayet, "Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino," *Nucl. Phys. B* **90** (1975) 104.

[39] R. Barbieri, S. Ferrara, and C. A. Savoy, "Gauge Models with Spontaneously Broken Local Supersymmetry," *Phys. Lett.* **B119** (1982) 343–352.

[40] L. J. Hall, J. D. Lykken, and S. Weinberg, "Supergravity as the Messenger of Supersymmetry Breaking," *Phys. Rev. D* **27** (1983) 2359.

[41] S. P. Martin, "A Supersymmetry primer," *Adv.Ser.Direct.High Energy Phys.* **21** (2010) 1–153, arXiv:hep-ph/9709356 [hep-ph].

[42] N. Polonsky, "Supersymmetry: Structure and phenomena. Extensions of the standard model," *Lect. Notes Phys.* **M68** (2001) 1–169, arXiv:hep-ph/0108236.

[43] M. Kramer, A. Kulesza, R. van der Leeuw, M. Mangano, S. Padhi, *et al.*, "Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV," arXiv:1206.2892 [hep-ph].

[44] R. Barbieri and D. Pappadopulo, "S-particles at their naturalness limits," *JHEP* **0910** (2009) 061, arXiv:0906.4546 [hep-ph].

[45] M. Papucci, J. T. Ruderman, and A. Weiler, "Natural SUSY Endures," *JHEP* **1209** (2012) 035, arXiv:1110.6926 [hep-ph].

[46] **LHC New Physics Working Group** Collaboration, D. Alves *et al.*, "Simplified Models for LHC New Physics Searches," *J.Phys.* **G39** (2012) 105005, arXiv:1105.2838 [hep-ph].

[47] J. Alwall, P. Schuster, and N. Toro, "Simplified Models for a First Characterization of New Physics at the LHC," *Phys.Rev.* **D79** (2009) 075020, arXiv:0810.3921 [hep-ph].

[48] **CMS** Collaboration, S. Chatrchyan *et al.*, "Interpretation of Searches for Supersymmetry with simplified Models," *Phys.Rev.* **D88** no. 5, (2013) 052017, arXiv:1301.2175 [hep-ex].

[49] S. Kraml, S. Kulkarni, U. Laa, A. Lessa, W. Magerl, *et al.*, "SModelS: a tool for interpreting simplified-model results from the LHC and its application to supersymmetry," *Eur.Phys.J.* **C74** (2014) 2868, arXiv:1312.4175 [hep-ph].

[50] S. Kraml, S. Kulkarni, U. Laa, A. Lessa, V. Magerl, *et al.*, "SModelS v1.0: a short user guide," arXiv:1412.1745 [hep-ph].

[51] M. Papucci, K. Sakurai, A. Weiler, and L. Zeune, "Fastlim: a fast LHC limit calculator," *Eur.Phys.J.* **C74** no. 11, (2014) 3163, arXiv:1402.0492 [hep-ph].

[52] J. S. Kim, D. Schmeier, J. Tattersall, and K. Rolbiecki, "A framework to create customised LHC analyses within CheckMATE," arXiv:1503.01123 [hep-ph].

[53] N. Arkani-Hamed, P. Schuster, N. Toro, J. Thaler, L.-T. Wang, *et al.*, "MARMOSET: The Path from LHC Data to the New Standard Model via On-Shell Effective Theories," arXiv:hep-ph/0703088 [HEP-PH].

[54] L. Evans and P. Bryant, "LHC Machine," *JINST* **3** (2008) S08001.

[55] O. S. Bruning and P. Collier, "Building a behemoth," *Nature* **448** (2007) 285–289.

[56] C. Lefevre, "LHC: the guide (English version)." http://cds.cern.ch/record/1165534, Feb, 2009.

[57] http://home.web.cern.ch/topics/large-hadron-collider.

[58] F. Marcastel, "CERN's Accelerator Complex.,". https://cds.cern.ch/record/1621583.

[59] S. Dailler, "Diagram of the beam dumps at LHC Point 6." http://cds.cern.ch/record/842348, Mar, 1997.

[60] R. Schmidt, R. Assmann, E. Carlier, B. Dehning, R. Denz, *et al.*, "Protection of the CERN Large Hadron Collider," *New J.Phys.* **8** (2006) 290.

[61] J.-L. Caron, "Cross section of LHC dipole." `https://cds.cern.ch/record/841539`, May, 1998.

[62] `http://www.lhc-closer.es`.

[63] **CMS** Collaboration, S. Chatrchyan *et al.*, "The CMS experiment at the CERN LHC," *JINST* **3** (2008) S08004.

[64] **CMS** Collaboration, G. L. Bayatian *et al.*, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software.* Technical Design Report CMS. CERN, Geneva, 2006. `http://cds.cern.ch/record/922757`.

[65] **CMS** Collaboration, G. Bayatian *et al.*, "CMS Physics: Technical Design Report Volume 2: Physics Performance," *J.Phys.* **G34** (2007) 995–1579.

[66] `http://cms.web.cern.ch/news/what-cms`.

[67] `http://cms.web.cern.ch/news/cms-detector-design`.

[68] CMS Document 4172-v2, `https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=4172`.

[69] **CMS** Collaboration, "Fish-eye view of yoke.." CMS Collection. `http://cds.cern.ch/record/710611`, Jun, 2002.

[70] **CMS** Collaboration, "First half of CMS inner tracker barrel." `http://cds.cern.ch/record/995912`, Oct, 2006.

[71] **CMS** Collaboration, "Images of the CMS ECAL Barrel (EB)." CMS Collection. `http://cds.cern.ch/record/1431477`, Nov, 2008.

[72] **CMS** Collaboration, "Images of the CMS HCAL Barrel (HB)." CMS Collection. `http://cds.cern.ch/record/1431485`, Nov, 2008.

[73] **CMS** Collaboration, "Images of CMS HCAL Forward Calorimeter (HF)." CMS Collection. `https://cds.cern.ch/record/1431489`, Nov, 2008.

[74] **CMS** Collaboration, "Muon Chambers." CMS Collection. `http://cds.cern.ch/record/1274467`, Dec, 2007.

[75] **CMS Trigger and Data Acquisition Group** Collaboration, W. Adam *et al.*, "The CMS high level trigger," *Eur.Phys.J.* **C46** (2006) 605–667, `arXiv:hep-ex/0512077 [hep-ex]`.

[76] **CMS** Collaboration, S. Cittolin, A. Rcz, and P. Sphicas, *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger.* Technical Design Report CMS. CERN, Geneva, 2002.

[77] P. Z. Skands, "QCD for Collider Physics," arXiv:1104.2863 [hep-ph].

[78] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, *et al.*, "Event generation with SHERPA 1.1," *JHEP* **0902** (2009) 007, arXiv:0811.4622 [hep-ph].

[79] R. Field, "Min-Bias and the Underlying Event at the LHC," arXiv:1202.0901 [hep-ph].

[80] J. M. Campbell, J. Huston, and W. Stirling, "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics," *Rept.Prog.Phys.* **70** (2007) 89, arXiv:hep-ph/0611148 [hep-ph].

[81] G. Salam, "Parton distribution functions." https://gsalam.web.cern.ch/gsalam/repository/talks/2009-Bautzen-lecture2.pdf. Lecture from the 2009 European School for High Energy Physics, in Bautzen, Germany.

[82] G. P. Salam, "Elements of QCD for hadron colliders," arXiv:1011.5131 [hep-ph].

[83] W. Tung, "Perturbative QCD and the parton structure of the nucleon,". http://www.physics.smu.edu/%7Eolness/cteqpp/tung2003/IntroPqcd.pdf.

[84] https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS12024.

[85] J. C. Collins, D. E. Soper, and G. F. Sterman, "Factorization of Hard Processes in QCD," *Adv.Ser.Direct.High Energy Phys.* **5** (1988) 1–91, arXiv:hep-ph/0409313 [hep-ph].

[86] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, *et al.*, "Parton distributions with LHC data," *Nucl.Phys.* **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[87] https://nnpdf.hepforge.org/.

[88] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, *et al.*, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP* **1407** (2014) 079, arXiv:1405.0301 [hep-ph].

[89] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP* **1106** (2011) 128, arXiv:1106.0522 [hep-ph].

[90] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[91] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche, *et al.*, "General-purpose event generators for LHC physics," *Phys.Rept.* **504** (2011) 145–233, arXiv:1101.2599 [hep-ph].

[92] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *JHEP* **0307** (2003) 001, arXiv:hep-ph/0206293 [hep-ph].

[93] A. Alloul, N. D. Christensen, C. Degrande, C. Duhr, and B. Fuks, "FeynRules 2.0 - A complete toolbox for tree-level phenomenology," *Comput.Phys.Commun.* **185** (2014) 2250–2300, arXiv:1310.1921 [hep-ph].

[94] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, *et al.*, "UFO - The Universal FeynRules Output," *Comput.Phys.Commun.* **183** (2012) 1201–1214, arXiv:1108.2040 [hep-ph].

[95] P. de Aquino, W. Link, F. Maltoni, O. Mattelaer, and T. Stelzer, "ALOHA: Automatic Libraries Of Helicity Amplitudes for Feynman Diagram Computations," *Comput.Phys.Commun.* **183** (2012) 2254–2263, arXiv:1108.2041 [hep-ph].

[96] F. Maltoni and T. Stelzer, "MadEvent: Automatic event generation with MadGraph," *JHEP* **0302** (2003) 027, arXiv:hep-ph/0208156 [hep-ph].

[97] J. Alwall, S. Hoche, F. Krauss, N. Lavesson, L. Lonnblad, *et al.*, "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions," *Eur.Phys.J.* **C53** (2008) 473–500, arXiv:0706.2569 [hep-ph].

[98] http://mlm.web.cern.ch/mlm/talks/lund-alpgen.pdf.

[99] J. Alwall, S. de Visscher, and F. Maltoni, "QCD radiation in the production of heavy colored particles at the LHC," *JHEP* **0902** (2009) 017, arXiv:0810.5350 [hep-ph].

[100] S. Banerjee, "Readiness of CMS Simulation Towards LHC Startup," *J.Phys.Conf.Ser.* **119** (2008) 032006.

[101] **CMS** Collaboration, S. Banerjee and M. Hildreth, "Validation and tuning of the CMS full simulation," *J.Phys.Conf.Ser.* **331** (2011) 032015.

[102] S. Banerjee, "CMS simulation software," *J.Phys.Conf.Ser.* **396** (2012) 022003.

[103] M. Dobbs and J. B. Hansen, "The HepMC C++ Monte Carlo event record for High Energy Physics," *Comput.Phys.Commun.* **134** (2001) 41–46.

[104] **GEANT4** Collaboration, S. Agostinelli *et al.*, "GEANT4: A simulation toolkit," *Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[105] **CMS** Collaboration, S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta, "The fast simulation of the CMS detector at LHC," *J.Phys.Conf.Ser.* **331** (2011) 032049.

[106] R. Rahmat, R. Kroeger, and A. Giammanco, "The fast simulation of the CMS experiment," *J.Phys.Conf.Ser.* **396** (2012) 062016.

[107] **CMS** Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," Tech. Rep. CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr, 2009.

[108] **CMS** Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV," Tech. Rep. CMS-PAS-PFT-10-002, CERN, Geneva, 2010.

[109] **CMS** Collaboration, S. Chatrchyan *et al.*, "Description and performance of track and primary-vertex reconstruction with the CMS tracker," *JINST* **9** no. 10, (2014) P10009, arXiv:1405.6569 [physics.ins-det].

[110] **CMS** Collaboration, V. Khachatryan *et al.*, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at sqrt(s)=8 TeV," arXiv:1502.02701 [physics.ins-det].

[111] **CMS** Collaboration, "Pileup Removal Algorithms," Tech. Rep. CMS-PAS-JME-14-001, CERN, Geneva, 2014. https://cds.cern.ch/record/1751454.

[112] G. P. Salam, "Towards Jetography," *Eur.Phys.J.* **C67** (2010) 637–686, arXiv:0906.1833 [hep-ph].

[113] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP* **0804** (2008) 063, arXiv:0802.1189 [hep-ph].

[114] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur.Phys.J.* **C72** (2012) 1896, arXiv:1111.6097 [hep-ph].

[115] https://twiki.cern.ch/twiki/bin/viewauth/CMS/IntroToJEC.

[116] **CMS** Collaboration, S. Chatrchyan *et al.*, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS," *JINST* **6** (2011) P11002, arXiv:1107.4277 [physics.ins-det].

[117] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys.Lett.* **B659** (2008) 119–126, arXiv:0707.1378 [hep-ph].

[118] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets," *JHEP* **0804** (2008) 005, arXiv:0802.1188 [hep-ph].

[119] https://twiki.cern.ch/twiki/bin/view/CMSPublic/MultipleConeSizes14.

[120] https://twiki.cern.ch/twiki/bin/viewauth/CMS/NewJECPlots13.

[121] **CMS** Collaboration, S. Chatrchyan *et al.*, "Identification of b-quark jets with the CMS experiment," *JINST* **8** (2013) P04013, arXiv:1211.4462 [hep-ex].

[122] **CMS** Collaboration, "Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events," tech. rep., CERN, Geneva, 2013. https://cds.cern.ch/record/1581306.

[123] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting," *J.Phys.* **G34** (2007) N343.

[124] https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagPerformanceOP.

[125] **CMS** Collaboration, "Scalar Top Quark Search with Jets and Missing Transverse Momentum in pp Collisions at sqrt(s) = 7 TeV," Tech. Rep. CMS AN2011/498, CERN, Geneva, 2012.

[126] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId.

[127] https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaCutBasedIdentification.

[128] **CMS** Collaboration, "Search for Direct Top Squark Pair Production in the Single Lepton Channel with Transverse Mass at 8 TeV," Tech. Rep. CMS AN-2013/089, CERN, Geneva, 2013.

[129] **CMS** Collaboration, V. Khachatryan *et al.*, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV," *JINST* **10** no. 02, (2015) P02006, arXiv:1411.0511 [physics.ins-det].

[130] https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters.

[131] https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupInformation.

[132] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for top-squark pair production in the single-lepton final state in pp collisions at $\sqrt{s} = 8$ TeV," *Eur.Phys.J.* **C73** (2013) 2677, arXiv:1308.1586 [hep-ex].

[133] https://twiki.cern.ch/twiki/bin/viewauth/CMS/SMST2ccMadgraph8TeV.

[134] **CMS** Collaboration, S. Chatrchyan *et al.*, "Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7$ TeV," *Eur.Phys.J.* **C73** (2013) 2339, arXiv:1211.2220 [hep-ex].

[135] https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting.

[136] **CMS** Collaboration, "Phenomenological MSSM interpretation of the CMS 7 and 8 TeV results," Tech. Rep. CMS-PAS-SUS-13-020, CERN, Geneva, 2014. https://cds.cern.ch/record/1693148.

[137] **ATLAS** Collaboration, G. Aad *et al.*, "Search for direct third-generation squark pair production in final states with missing transverse momentum and two b-jets in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector," *JHEP* **1310** (2013) 189, arXiv:1308.2631 [hep-ex].

[138] **ATLAS** Collaboration, G. Aad *et al.*, "Search for direct top-squark pair production in final states with two leptons in pp collisions at $\sqrt{s} =$ 8TeV with the ATLAS detector," *JHEP* **1406** (2014) 124, arXiv:1403.4853 [hep-ex].

[139] **ATLAS** Collaboration, G. Aad *et al.*, "Search for direct pair production of the top squark in all-hadronic final states in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *JHEP* **1409** (2014) 015, arXiv:1406.1122 [hep-ex].

[140] **ATLAS** Collaboration, G. Aad *et al.*, "Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector," *Eur.Phys.J.* **C74** no. 10, (2014) 3109, arXiv:1406.5375 [hep-ex].

[141] **ATLAS** Collaboration, G. Aad *et al.*, "Search for top squark pair production in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector," arXiv:1407.0583 [hep-ex]. Submitted for publication.

[142] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for top squark and higgsino production using diphoton Higgs boson decays," *Phys.Rev.Lett.* **112** (2014) 161802, arXiv:1312.3310 [hep-ex].

[143] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for top-squark pairs decaying into Higgs or Z bosons in pp collisions at $\sqrt{s}$=8 TeV," *Phys.Lett.* **B736** (2014) 371–397, arXiv:1405.3886 [hep-ex].

[144] **CMS** Collaboration, "Search for top squarks decaying to a charm quark and a neutralino in events with a jet and missing transverse momentum," Tech. Rep. CMS-PAS-SUS-13-009, CERN, Geneva, 2014. https://cds.cern.ch/record/1644584.

[145] **ATLAS** Collaboration, G. Aad *et al.*, "Search for pair-produced third-generation squarks decaying via charm quarks or in compressed supersymmetric scenarios in *pp* collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Phys.Rev.* **D90** (2014) 052008, arXiv:1407.0608 [hep-ex].

[146] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure," *Phys.Rev.* **D80** (2009) 051501, arXiv:0903.5081 [hep-ph].

[147] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches," *Phys.Rev.* **D81** (2010) 094023, arXiv:0912.0033 [hep-ph].

[148] **CMS** Collaboration, S. Chatrchyan *et al.*, "Studies of jet mass in dijet and W/Z + jet events," *JHEP* **1305** (2013) 090, arXiv:1303.4811 [hep-ex].

[149] J. Thaler and K. Van Tilburg, "Identifying Boosted Objects with N-subjettiness," *JHEP* **1103** (2011) 015, arXiv:1011.2268 [hep-ph].

[150] C. Lester and D. Summers, "Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders," *Phys.Lett.* **B463** (1999) 99–103, arXiv:hep-ph/9906349 [hep-ph].

[151] A. Barr, C. Lester, and P. Stephens, "m(T2): The Truth behind the glamour," *J.Phys.* **G29** (2003) 2343–2363, arXiv:hep-ph/0304226 [hep-ph].

[152] L. Randall and D. Tucker-Smith, "Dijet Searches for Supersymmetry at the LHC," *Phys.Rev.Lett.* **101** (2008) 221803, arXiv:0806.1049 [hep-ph].

[153] G. Polesello and D. R. Tovey, "Supersymmetric particle mass measurement with the boost-corrected contransverse mass," *JHEP* **1003** (2010) 030, arXiv:0910.0174 [hep-ph].

[154] Y. Bai, H.-C. Cheng, J. Gallicchio, and J. Gu, "Stop the Top Background of the Stop Search," *JHEP* **1207** (2012) 110, arXiv:1203.4813 [hep-ph].

[155] C. Rogan, "Kinematical variables towards new dynamics at the LHC," arXiv:1006.2727 [hep-ph].

[156] C. S. Rogan and M. Spiropulu, *Searches for new symmetries in pp collisions with the razor kinematic variables at* $\sqrt{s} = 7$ *TeV*. PhD thesis, Caltech, 2013. presented 12 Feb 2013.

[157] **CMS** Collaboration, S. Chatrchyan *et al.*, "Inclusive search for squarks and gluinos in *pp* collisions at $\sqrt{s} = 7$ TeV," *Phys.Rev.* **D85** (2012) 012004, arXiv:1107.1279 [hep-ex].

[158] **CMS** Collaboration, S. Chatrchyan *et al.*, "Search for supersymmetry with razor variables in pp collisions at $\sqrt{s}$=7TeV," *Phys.Rev.* **D90** no. 11, (2014) 112001, arXiv:1405.3961 [hep-ex].

[159] **CMS** Collaboration, V. Khachatryan *et al.*, "Identification techniques for highly boosted W bosons that decay into hadrons," *JHEP* **1412** (2014) 017, arXiv:1410.4227 [hep-ex].

[160] D. Krohn, J. Thaler, and L.-T. Wang, "Jet Trimming," *JHEP* **1002** (2010) 084, arXiv:0912.1342 [hep-ph].

[161] J. Gallicchio and M. D. Schwartz, "Seeing in Color: Jet Superstructure," *Phys.Rev.Lett.* **105** (2010) 022001, arXiv:1001.5027 [hep-ph].

[162] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC," *Phys.Rev.Lett.* **100** (2008) 242001, arXiv:0802.2470 [hep-ph].

[163] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, "Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks," *Phys.Rev.Lett.* **101** (2008) 142001, arXiv:0806.0848 [hep-ph].

[164] **CMS** Collaboration, "Identifying Hadronically Decaying Vector Bosons Merged into a Single Jet," Tech. Rep. CMS-PAS-JME-13-006, CERN, Geneva, 2013. https://cds.cern.ch/record/1577417.

[165] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, "Better jet clustering algorithms," *JHEP* **9708** (1997) 001, arXiv:hep-ph/9707323 [hep-ph].

[166] **CMS** Collaboration, "Boosted Top Jet Tagging at CMS," Tech. Rep. CMS-PAS-JME-13-007, CERN, Geneva, 2014.

[167] **CMS** Collaboration, "Search for heavy resonances in the W/Z-tagged dijet mass spectrum in pp collisions at 8 TeV," Tech. Rep. CMS AN2012/393, CERN, Geneva, 2012.

[168] S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions," *Phys.Rev.* **D48** (1993) 3160–3166, arXiv:hep-ph/9305266 [hep-ph].

[169] S. Catani, Y. L. Dokshitzer, M. Seymour, and B. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions," *Nucl.Phys.* **B406** (1993) 187–224.

[170] "N-subjettiness implementation in fastjet-contrib." http://fastjet.hepforge.org/svn/contrib/contribs/Nsubjettiness/tags/2.1.0/README. Accessed: 2015/01/19.

[171] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at $\sqrt{s} = 8$ TeV," *JHEP* **1408** (2014) 173, arXiv:1405.1994 [hep-ex].

[172] **CMS** Collaboration, V. Khachatryan *et al.*, "Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at $\sqrt{s} = 8$ TeV," *JHEP* **1408** (2014) 174, arXiv:1405.3447 [hep-ex].

[173] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP* **0207** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[174] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions," *JHEP* **0909** (2009) 111, arXiv:0907.4076 [hep-ph].

[175] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method," *Eur.Phys.J.* **C71** (2011) 1547, arXiv:1009.2450 [hep-ph].

[176] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, *et al.*, "New parton distributions for collider physics," *Phys.Rev.* **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[177] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP* **0206** (2002) 029, arXiv:hep-ph/0204244 [hep-ph].

[178] **CMS** Collaboration, S. Chatrchyan *et al.*, "Study of the underlying event at forward rapidity in pp collisions at $\sqrt{s} = 0.9$, 2.76, and 7 TeV," *JHEP* **1304** (2013) 072, arXiv:1302.2394 [hep-ex].

[179] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, *et al.*, "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)," *JHEP* **0101** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[180] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, *et al.*, "HERWIG 6.5 release note," arXiv:hep-ph/0210213 [hep-ph].

[181] **CMS** Collaboration, S. Chatrchyan *et al.*, "Inclusive Search for Supersymmetry Using Razor Variables in *pp* Collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev. Lett.* **111** (Aug, 2013) 081802, arXiv:1212.6961 [hep-ex].

[182] **CMS and ATLAS** Collaboration, "Procedure for the lhc higgs boson search combination in summer 2011," Tech. Rep. CMS-NOTE-2011-005, CERN, Geneva, Aug, 2011. http://cds.cern.ch/record/1379837.

[183] A. Martin, W. Stirling, R. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur.Phys.J.* **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[184] **NNPDF** Collaboration, R. D. Ball *et al.*, "Parton distributions with QED corrections," *Nucl.Phys.* **B877** (2013) 290–320, arXiv:1308.0598 [hep-ph].

[185] S. Alekhin, S. Alioli, R. D. Ball, V. Bertone, J. Blumlein, *et al.*, "The PDF4LHC Working Group Interim Report," arXiv:1101.0536 [hep-ph].

[186] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, *et al.*, "The PDF4LHC Working Group Interim Recommendations," arXiv:1101.0538 [hep-ph].

[187] J. Butterworth, G. Dissertori, S. Dittmaier, D. de Florian, N. Glover, *et al.*, ""New access to PDF data via LHAPDF6" in Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report," arXiv:1405.1067 [hep-ph].

[188] https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagSFMethods.

[189] https://twiki.cern.ch/twiki/bin/view/Main/EGammaScaleFactors2012#2012_8_TeV_Jan22_Re_recoed_data.